

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:11:13 ; Search time 126 Seconds  
(without alignments)  
1004.613 Million cell updates/sec

Title: US-10-650-369-22  
Perfect score: 2278  
Sequence: 1 MKKIKITGILLLAVILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

## ALIGNMENTS

Abb55601 Lactococc  
Aau35254 Enterococ  
Abu29370 Protein e  
Abb53868 Lactococc  
Adc95142 E. faeciu  
Abu25342 Protein e  
Abb48810 Listeria  
Abu32568 Protein e  
Abu32559 Protein e  
Abu37077 Protein e  
Aag83058 S. epider  
Abu42966 Protein e  
Abp40668 Staphyloc  
Aau36947 Staphyloc  
Abu34122 Pathogen  
Abu19195 Protein e  
Abu16073 Protein e

RESULT 1  
AAM50664  
ID AAM50664 standard; protein; 448 AA.  
XX  
AC AAM50664;  
XX  
DT 29-AUG-2003 (revised)  
DT 08-APR-2002 (first entry)

Streptococcus GapC multiepitope fusion PolyGap4.  
DE  
XX  
KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;  
KW immunisation; mastitis; therapy.  
XX  
OS Streptococcus dysgalactiae.  
OS Streptococcus agalactiae.  
OS Streptococcus parauberis.  
OS Chimeric.  
XX  
WO200196379-A2.

XX  
PD 20-DEC-2001.  
XX  
PF 11-JUN-2001; 2001WO-CA000836.  
XX  
PR 12-JUN-2000; 2000US-0211247P.  
XX  
(USSA-) UNIV SASKATCHEWAN.  
PA  
XX  
PI Potter AA, Perez-Casal J, Fontaine M;  
XX  
WPI; 2002-098051/13.  
N-PSDB; ABA91327.

XX  
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic  
determinant of Streptococcus dysgalactiae, *S. agalactiae*, *S. uberis*, *S.*  
paruberis, or *S. iniae* GapC protein useful for treating mastitis in  
vertebrates.  
XX  
PS Claim 9; Fig 6A-C; 116pp; English.

CC The present sequence is that of a novel multiple epitope fusion protein,  
designated PolyGap4, comprising the entire amino acid sequence of the  
Streptococcus dysgalactiae GapC plasmin binding protein in addition to  
unique amino acid sequences from the Streptococcus parauberis and  
Streptococcus agalactiae GapC proteins. The multiple epitope protein is  
produced in host cells transformed with an expression vector comprising a  
chimeric gene (see ABA91327) encoding the protein. PolyGal4 is an example

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2278	100.0	448	5	AAM50664		Aam50664 Streptoco
2	1656.5	72.7	336	5	AAM50665		Aam50665 Streptoco
3	1656.5	72.7	336	5	AAM50639		Aam50639 Streptoco
4	1655.5	72.7	336	5	ABP29960		Abp29960 Streptoco
5	1655.5	72.7	336	6	ABU46455		Abu46455 Protein e
6	1655.5	72.7	345	5	ABP29106		Abp29106 Streptoco
7	1652.5	72.5	336	2	AAR56486		Aar56486 Plasmin r
8	1652.5	72.5	336	4	AAY85681		Aay85681 Streptoco
9	1564.5	68.7	336	5	ABP30758		Abp30758 Streptoco
10	1559.5	68.5	336	5	AAM50667		Aam50667 Streptoco
11	1559.5	68.5	336	5	AAM50641		Aam50641 Streptoco
12	1557.5	68.4	336	5	AAM50666		Aam50666 Streptoco
13	1557.5	68.4	336	5	AAM50640		Aam50640 Streptoco
14	1535	67.4	335	6	ABU02516		Abu02516 S. pneumo
15	1535	67.4	359	4	AUU37576		Aau37576 Streptoco
16	1535	67.4	359	6	ABU46262		Abu46262 Protein e
17	1534.5	67.4	336	5	AAM50669		Aam50669 Streptoco
18	1534.5	67.4	336	5	AAM50643		Aam50643 Streptoco
19	1534	67.3	359	4	AUU38000		Aau38000 Streptoco
20	1530.5	67.2	336	5	AAM50668		Aam50668 Streptoco
21	1530.5	67.2	336	5	AAM50642		Aam50642 Streptoco
22	1521	66.8	337	6	ABU44330		Abu44330 Protein e
23	1491	65.5	333	2	AAM50509		Aam50509 Streptoco
24	1491	65.5	333	5	ABP54583		Abp54583 S. pneumo
25	1491	65.5	333	7	ADG45135		Adc45135 S. pneumo

CC of novel GapC multiple epitope fusion proteins of the invention that  
 CC comprise epitopes from 1 or more of *S. dysgalactiae*, *S. agalactiae*, *S.  
 CC parauberis*, *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-  
 CC 69). The multiple epitope fusion proteins are used in claimed vaccines  
 CC for treating or preventing a bacterial infection in a vertebrate, especially a  
 CC streptococcal infection, and particularly mastitis. They are  
 CC also used in claimed methods of detecting *Streptococcus* antibodies. The  
 CC multiple epitope protein is capable of eliciting broad immunity against a  
 CC variety of streptococcal infections while minimising the number of  
 CC antigens present in the final formulation and concomitantly reducing  
 CC production costs. (Updated on 29-AUG-2003 to standardise OS field)

XX SQ Sequence 448 AA;

Query Match 100.0%; Score 2278; DB 5; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-180;  
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKITGILLAVIILSACQANYGSGMMVVKVTINGFGRIGRLAFRRIQVNVEGVETRIN 60  
 Db 1 MKKITGILLAVIILSACQANYGSGMMVVKVTINGFGRIGRLAFRRIQVNVEGVETRIN 60  
 QY 61 DLTDPMMLAHLKYDTTQGRFDTVEKVEGGFEVNGNFIKVSAERDPENIDWATDGEIV 120  
 Db 61 DLTDPMMLAHLKYDTTQGRFDTVEKVEGGFEVNGNFIKVSAERDPENIDWATDGEIV 120  
 QY 121 LEALEGTVVKDGGFDVNGKFKIVSAEKDPEQIDWATDGEIVLEIDGTEVKEGGFEVN 180  
 Db 121 LEALEGTVVKDGGFDVNGKFKIVSAEKDPEQIDWATDGEIVLEIDGTEVKEGGFEVN 180  
 QY 181 GQFVKVSAEREPAKIDWATDGEIVLEATSEFAKEAAEKHLHANGAKKKVITAPGGNDV 240  
 Db 181 GQFVKVSAEREPAKIDWATDGEIVLEATSEFAKEAAEKHLHANGAKKKVITAPGGNDV 240  
 QY 241 KTVVFNTNHDILDGTETVVISGASCTNNCLAPMAKALHDAFGIQKGLMTTHAYTGDQML 300  
 Db 241 KTVVFNTNHDILDGTETVVISGASCTNNCLAPMAKALHDAFGIQKGLMTTHAYTGDQML 300  
 QY 301 DGPHRGDDLRRARAAGAANTVPNSTGAAKAIGLVIPELNGKLDGAQQRVPVPTGSVTELVV 360  
 Db 301 DGPHRGDDLRRARAAGAANTVPNSTGAAKAIGLVIPELNGKLDGAQQRVPVPTGSVTELVV 360  
 QY 361 TLDKNTVSDEINAAMKAASNSDSFGYTEDPIVSSDTIVGVSYGSLFDATQTKVMEVDGSQLV 420  
 Db 361 TLDKNTVSDEINAAMKAASNSDSFGYTEDPIVSSDTIVGVSYGSLFDATQTKVMEVDGSQLV 420  
 QY 421 KVVSWDNEMSYTAQLVRTLEYFAKIAK 448  
 Db 421 KVVSWDNEMSYTAQLVRTLEYFAKIAK 448

RESULT 2  
 AAM50665 standard; protein; 336 AA.  
 XX AC AAM50665;  
 XX DT 08-APR-2002 (first entry)

XX DE Streptococcus dysgalactiae gapC plasmin binding protein DysGapC.  
 XX KW DysGapC; GapC; Plasmin binding protein; epitope; infection; vaccine;  
 KW immunisation; mastitis; therapy.  
 XX OS Streptococcus dysgalactiae.  
 XX PN WO200196379-A2.  
 XX PD 20-DEC-2001.  
 XX PP 11-JUN-2001; 2001WO-CA0000836.  
 XX PR 12-JUN-2000; 2000US-0211247P.

RESULT 3  
 AAM50639  
 XX PA (UUSA-) UNIV SASKATCHEWAN.  
 XX PA Potter AA, Perez-Casal J, Fontaine M;  
 XX PI WPI; 2002-098051/13.  
 XX DR N-PSDB; ABA91328.  
 XX DR Novel GapC multiple epitope fusion polypeptide comprising antigenic  
 PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.  
 PT parauberis*, or *S. iniae* GapC protein useful for treating mastitis in  
 PT vertebrates.  
 XX PS Claim 8; Fig 1A-B; 116pp; English.

XX The present sequence is that of the GapC plasmin binding protein,  
 CC DysGapC, of *Streptococcus dysgalactiae* ATCC 43078, an isolate from a case  
 CC of bovine mastitis. The invention relates to novel GapC multiple epitope  
 CC fusion proteins that comprise epitopes from 1 or more of *Streptococcus*  
 CC *dysgalactiae*, *Streptococcus agalactiae*, *Streptococcus parauberis*,  
 CC *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). A claimed  
 CC example is Polygap4 (see AAM50664). Expression vectors and host cells for  
 CC production of the multiple epitope fusion proteins are provided. The  
 CC multiple epitope proteins are used in claimed vaccines for treating or  
 CC preventing a bacterial infection in a vertebrate, especially a  
 CC streptococcal infection, and particularly mastitis. They are also used in  
 CC claimed methods of detecting *Streptococcus* antibodies. The multiple  
 CC epitope proteins are capable of eliciting broad immunity against a  
 CC variety of streptococcal infections while minimising the number of  
 CC antigens present in the final formulation and concomitantly reducing  
 CC production costs  
 XX SQ Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;  
 Best Local Similarity 79.6%; Pred. No. 4.9e-129;  
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVGINGFGRIGRLAFRRIQVNVEGVETRINDLTQGRFDTGTVEV 87  
 Db 1 MVVKVGINGFGRIGRLAFRRIQVNVEGVETRINDLTQGRFDTGTVEV 60  
 QY 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGEIVLEALEGTVVKDGGFDVNGKFKIYVSAE 147  
 Db 61 KEGGFEVNGNFIKVSAAERDPENIDWATDGEIVLEALEGTVVKDGGFDVNGKFKIYVSAE 81  
 QY 148 KDPEQIDWATDGEIVLEIDGTEVKEGGFEVNGQFFVKAEREPAKIDWATDGEIVLE 207  
 Db 82 -----  
 QY 208 ATSFPEAKKEAAEKHLHANGAKKKVITAPGGNDVKTIVSGASCTTN 267  
 Db 96 ATGFFAKREAAEKHLHANGAKKKVITAPGGNDVKTIVSGASCTTN 27  
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPHRGDLRARRAGAANTVPNSTGAA 327  
 Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPHRGDLRARRAGAANTVPNSTGAA 215  
 QY 328 KAIGLVIPELNGKLDGAQQRVPVPTGSVTELVVLDKNTVSDEINAAMKAASNSDFGYTE 387  
 Db 216 KAIGLVIPELNGKLDGAQQRVPVPTGSVTELVVLDKNTVSDEINAAMKAASNSDFGYTE 275  
 QY 388 DPIVSSDIVGVSYGSLFDATQTKYMEVDGSQLVKVSAAERDPENIDWATDGEIVLE 447  
 Db 276 DPIVSSDIVGVSYGSLFDATQTKYMEVDGSQLVKVSAAERDPENIDWATDGEIVLE 335  
 QY 448 K 448  
 Db 336 K 336

ID	AAM50639	standard; protein; 336 AA.	QY	328 KAIGLVIPELNGKLDAQRVVPPTGSVTELWVTLDKNVSVDEINAAMKAASNDNSFGYTE	387
XX	AAM50639;		Db	216 KAIGLVIPELNGKLDAQRVVPPTGSVTELWVTLDKNVSVDEINAAMKAASNDNSFGYTE	275
AC					
XX	04-APR-2002	(First entry)	QY	388 DPVSSDIVGVSYGSLSFDATQTKVMEVDGSQQLVQVSWYDNEMSYTAQLVRLTLEYFAKIA	447
DT		Streptococcus dysgalactiae gapC plasmin binding protein.	Db	276 DPVSSDIVGVSYGSLSFDATQTKVMEVDGSQQLVQVSWYDNEMSYTAQLVRLTLEYFAKIA	335
XX					
DE					
XX	GapC; plasmin-binding protein; DysgalpC; infection; mastitis; vaccine;		QY	448 K 448	
KW	KW		Db	336 K 336	
XX	diagnosis; therapy.				
OS	Streptococcus dysgalactiae.				
XX					
PN	WO200196381-A2.				
XX					
PD	20-DEC-2001.				
XX					
PF	11-JUN-2001; 2001WO-CA000838.				
XX					
PR	12-JUN-2000; 2000US-0211022P.				
XX					
PA	(UYSA-) UNIV SASKATCHEWAN.				
XX					
PI	Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;				
XX					
PT	Novel isolated GapC protein of Streptococcus dysgalactiae, <i>S. agalactiae</i> , <i>S. uberis</i> , <i>S. paruberis</i> , or <i>S. iniae</i> , useful as vaccine component for treating streptococcal infection which causes mastitis in vertebrates.				
PT					
PT					
XX					
DR	WPI; 2002-130725/17.				
DR	N-PSDB; ABA91248.				
PS	Claim 1(a): Fig 1A-B; 107pp; English.				
XX					
CC	The present sequence is that of the GapC plasmin binding protein (DysGalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate from bovine mastitis. It is encoded by the GapC gene, given in ABA91248.				
CC	GapC protein, which has no signal sequence or membrane anchor domain, is capable of eliciting an immune response in a vertebrate. The invention provides the GapC genes and proteins of 5 streptococcus species, as well as recombinant vectors, host cells and vaccine compositions comprising GapC polynucleotides or proteins. The vaccines are used to treat or prevent a bacterial infection, especially a streptococcal infection, and mastitis in particular (claimed). GapC proteins are also used in claimed methods for detecting GapC antibodies, and to raise antibodies that are used in claimed methods for detecting GapC proteins				
XX					
XX	Sequence 336 AA;				
CC	Query Match 72.7%; Score 1656.5; DB 5; Length 336;				
CC	Best Local Similarity 79.6%; Pred. No. 4.9e-129;				
CC	Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;				
QY	28 MVVKVINGFGRIGRLAFRRIQNVEGVTRINDLTDPNMLAHLKDYDTTQGRFDGTVE	87			
Db	1 MVVKVINGFGRIGRLAFRRIQNVEGVTRINDLTDPNMLAHLKDYDTTQGRFDGTVE	60			
QY	88 KEGGFEVNGNFIKVAERDOPENIDWATDGEVLEALELEGTVKDFDNGKFLIKVSAE	147			
Db	61 KEGGFEVNGNFIKVAERDPE-----	81			
QY	148 KDPEQIDWATDGEVIEVLIDGTVEVKEGGFEVNGQFKVSAEREPANIDWATDGEVIEVL	207			
Db	82 -----	-----			
QY	208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVVKTVFNTNHDILDGTETVIGASCTTN	267			
Db	96 ATGFFPAKKEAAEKHLHANGAKKVVITAPGGNDVVKTVFNTNHDILDGTETVIGASCTTN	155			
QY	268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGPHRGDDLREARAGAANIVPNSTGAA	327			
Db	156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGPHRGDDLREARAGAANIVPNSTGAA	215			
XX	Sequence 336 AA;				
SQ					

Query Match 72.7%; Score 1655.5; DB 5; Length 336;  
 Best Local Similarity 79.3%; Pred. No. 5.9e-129;  
 Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

PT for homologous nucleic acids required for cellular proliferation to  
 isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 74379; 1766pp; English.

PT  
 PT  
 XX  
 PS  
 XX

CC The invention relates to an isolated nucleic acid comprising any one of  
 the 6213 antisense sequences given in the specification where expression  
 of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 proliferation or the activity of a gene in an operon required for  
 proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs; or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

XX

QY 28 MVVKVINGFGRIGRLAERRRIONVEGVETRINDLTDPMMLAHLKYDTTQGRFDTVEV 87  
 Db 1 MVVKVINGFGRIGRLAERRRIONIEGVETRINDLTDPMMLAHLKYDTTQGRFDTVEV 60

QY 88 KEGGFEVNGNFIKVAERDPENIDWATDGVEIVLEALEGTVEVKGDFDNGKFKIKVSAE 147  
 Db 61 KEGGFEVNGNFIKVAERDP----- 81

QY 148 KDPEQIDWATDGVEIVLEIDGTVEVKGDFEVNGQFVKVSAEREPANIDWATDGVEIVLE 207  
 Db 82 -----NIDWATDGVEIVLE 95

QY 208 ATSEFAKKEAAEKLHANGAKKVKVITAPGGNDVKTVVVENTNHDILDGTEVIGSACSTTN 267  
 Db 96 ATGFFAKKEAAEKLHANGAKKVKVITAPGGNDVKTVVVENTNHDILDGTEVIGSACSTTN 155

QY 268 CLAPMAKALHDQAFGIQKGLMTTIHAYTQDMILDQPHRGDLARRAGAANIVPNSTGAA 327  
 Db 156 CLAPMAKALHDQAFGIQKGLMTTIHAYTQDMILDQPHRGDLARRAGAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRYVPVPTGSVTELVVTLDKNNVSVDINAAMKAASNDSFGYTE 387  
 Db 216 KAIGLVIPELNGKLDGAAQRYVPVPTGSVTELVVTLDKNNVSVDINAAMKAASNDSFGYTE 275

QY 388 DPTVSSDIVGVSYGSLFDATQTKYMEDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 447  
 Db 276 DPTVSSDIVGVSYGSLFDATQTKYMEDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448  
 Db 336 K 336

XX

SQ Sequence 336 AA;

Query Match 72.7%; Score 1655.5; DB 6; Length 336;  
 Best Local Similarity 79.3%; Pred. No. 5.9e-129;  
 Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAERRRIONVEGVETRINDLTDPMMLAHLKYDTTQGRFDTVEV 87  
 Db 1 MVVKVINGFGRIGRLAERRRIONIEGVETRINDLTDPMMLAHLKYDTTQGRFDTVEV 60

QY 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVEVKGDFDNGKFKIKVSAE 147  
 Db 61 KEGGFEVNGNFIKVSAAERDP----- 81

QY 148 KDPEQIDWATDGVEIVLEIDGTVEVKGDFEVNGQFVKVSAEREPANIDWATDGVEIVLE 207  
 Db 82 -----NIDWATDGVEIVLE 95

QY 208 ATSEFAKKEAAEKLHANGAKKVKVITAPGGNDVKTVVVENTNHDILDGTEVIGSACSTTN 267  
 Db 96 ATGFFAKKEAAEKLHANGAKKVKVITAPGGNDVKTVVVENTNHDILDGTEVIGSACSTTN 155

QY 268 CLAPMAKALHDQAFGIQKGLMTTIHAYTQDMILDQPHRGDLARRAGAANIVPNSTGAA 327  
 Db 156 CLAPMAKALHDQAFGIQKGLMTTIHAYTQDMILDQPHRGDLARRAGAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRYVPVPTGSVTELVVTLDKNNVSVDINAAMKAASNDSFGYTE 387  
 Db 216 KAIGLVIPELNGKLDGAAQRYVPVPTGSVTELVVTLDKNNVSVDINAAMKAASNDSFGYTE 275

QY 388 DPTVSSDIVGVSYGSLFDATQTKYMEDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 447  
 Db 276 DPTVSSDIVGVSYGSLFDATQTKYMEDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448

PT New antisense nucleic acids, useful for identifying proteins or screening



plasminogen activator (PA) so allows a reduction in dose, and thus lowers the risk of bleeding, and may prevent reocclusion of blood vessels. The protein may be coupled to a fibrin-specific monoclonal antibody to provide targetting to clots. The plasmin receptor may be useful in human or veterinary medicine, for treatment of thrombosis and pulmonary embolism, and for solubilising clots in catheters or shunts. (Updated on 25-MAR-2003 to correct PF field.)

Von Mering G, Broder C, Lottenberg R, Boyle MDP;  
WPI; 2001-006210/01.  
N-PSDB; AAC66140.

Raising an immune response in a mammal, especially for vaccination against group A streptococcal infections, comprises administering an isolated streptococcal plasmin receptor protein

Job 2Y		82		- - - - -		207		KDPEQIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVEIVL		207	
QY	208	ATSFFAKKEAAEKHLHANGAKKVITAPEGGNDVKTIVVENTNHDILDGTETVISGASCTTN	267	QY	28	MVVKVGINGFGRIGRLAFRRRIQNEGVTRINDLTDPNMLAHLKYDTTQGRFEDGTVEV	87	Score 1652.5; DB 4; Length 336;	72.5%; Score 1652.5; DB 4; Length 336;	Best Local Similarity 79.1%; Pred. No. 1e-128; Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1	79.1%; Pred. No. 1e-128; Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1
Db	96	ATGFFAKKEAAEKHLHANGAKKVITAPEGGNDVKTIVVENTNHDILDGTETVISGASCTTN	267	Db	1	MVVKVGINGFGRIGRLAFRRRIQNEGVTRINDLTDPNMLAHLKYDTTQGRFEDGTVEV	60				
QY	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAGAANIVPNSTGAA	327	QY	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVEVKDGGFEDVNGKFIKVSAA	147				
Db	156	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAGAANIVPNSTGAA	327	Db	61	KEGGFEVNGNFIKVSAAERDP	215				
QY	328	KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAKAAASNDSEFGYTE	387	QY	148	KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVEIVL	207				
Db	216	KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAKAAASNDSEFGYTE	387	Db	82						
QY	388	DPIVSSDITVGVSYGSLSLFDATQTKVMEVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA	447	QY	208	ATSFPEAKKEAAEKHLHANGAKKVITAPEGGNDVKTIVVENTNHDILDGTETVISGASCTTN	267				
Db	276	DPIVSSDITVGVSYGSLSLFDATQTKVMEVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA	335	Db	96	ATGFFPEAKKEAAEKHLHANGAKKVITAPEGGNDVKTIVVENTNHDILDGTETVISGASCTTN	155				
QY	448	K 448		QY	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAGAANIVPNSTGAA	327				
Db	336	K 336		Db	156	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAGAANIVPNSTGAA	215				
QY	328	KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAKAAASNDSEFGYTE	387	QY	328	KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAKAAASNDSEFGYTE	387				
Db	216	KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAKAAASNDSEFGYTE	387	Db	216	KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAKAAASNDSEFGYTE	387				
RESULT 8											
QY	AY85681			QY	388	DPIVSSDITVGVSYGSLSLFDATQTKVMEVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA	447				
Db	AY85681	standard; protein; 336 AA.		Db	276	DPIVSSDITVGVSYGSLSLFDATQTKVMEVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA	335				
Streptococcal plasmin receptor amino acid sequence.											
QY	AY85681;			QY	448	K 448					
Db	US6136323-A.			Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	13-FEB-2001	(first entry)		QY	448	K 448					
Db	02-JUL-2002	(first entry)		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	11-JUL-1994;	94US-00273247.		QY	448	K 448					
Db	24-OCT-2000.			Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448					
Db	16-MAY-1990;	90US-00524411.		Db	336	K 336					
QY	10-AUG-1992;	92US-00928462.		QY	448	K 448					
Db	11-JUL-1994;	94US-00273247.		Db	336	K 336					
Streptococcal plasmin receptor amino acid sequence.											
QY	29-MAR-1989;	89US-00330849.		QY	448	K 448				</td	



1 MVVKVINGFGRIGLAFERRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTEV 60  
 88 KEGGFEVNGNFIKVAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVS 147  
 61 KDGGEVNGNFIKVAERDP----- 81  
 148 KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANTDWATDGVEIV 207  
 82 -----NIDWATDGVEIV 95  
 DB  
 QY 208 ATSFIAKKEAEKHLHANGAKKKVITAPGGNDVKTVVENTNHDILDGTETVVISGASCTTN 267  
 96 ATGFFAKKAAAEKHLHANGAKKKVITAPGGDDVKTVVENTNHDILDGTETVVISGASCTTN 155  
 DB  
 QY 268 CLAPMAKALHDAGIQLQGLMTTIHAYTQDMILDGPHRGGDLRRARAAGAANTIPVN 327  
 156 CLAPMAKALQDNGFVQQLGLMTTIHAYTQDMILDGPHRGGDLRRARAAGAANTIPVN 215  
 QY 328 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASND 387  
 216 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVAVLEKETSVEINAAMKAASND 275  
 QY 388 DPTVSSDIVGVSYGSLSFDATOTIKVMEVDGSQLVKVVSWYDNEMSYTAQLV 447  
 276 DPTVSSDIVGMAYSLSLFDATQTKVQTVDGNQLVKVVSWYDNEMSYTAQLV 335  
 QY 448 K 448  
 DB 336 K 336  
 RESULT 11  
 ID AAM50641 standard; protein; 336 AA.  
 XX  
 AC AAM50641;  
 XX  
 DT 04-APR-2002 (first entry)  
 XX  
 DE Streptococcus uberis gapC plasmin binding protein.  
 XX  
 KW GapC; plasmin-binding protein; ÜberGapC; infection; mastitis; vaccine;  
 KW diagnosis; therapy.  
 XX  
 OS Streptococcus uberis.  
 XX  
 PN WO200196381-A2.  
 XX  
 PD 20-DEC-2001.  
 XX  
 PF 11-JUN-2001; 2001WO-CA0000838.  
 XX  
 PR 12-JUN-2000; 2000US-0211022P.  
 XX  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 XX  
 PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;  
 XX  
 DR 2002-130725/17.  
 XX  
 DR N-PSDB; ABA91250.  
 XX  
 PS Claim 1(c); Fig 3A-B; 107pp; English.  
 XX  
 PT Novel isolated GapC protein of Streptococcus dysgalactiae, *S. agalactiae*,  
 PT *S. uberis*, *S. paruberis*, or *S. iniae*, useful as vaccine component for  
 PT treating streptococcal infection which causes mastitis in vertebrates.  
 XX  
 PS The present sequence is that of the GapC plasmin binding protein  
 PT (*ÜberGapC*) of *Streptococcus uberis* ATCC 9927, encoded by the GapC gene  
 PT given in ABA91250. GapC, which has no signal sequence or membrane anchor  
 PT domain, is capable of eliciting an immune response in a vertebrate. The  
 PT invention provides the GapC genes and proteins of 5 *Streptococcus*  
 CC species, as well as recombinant vectors, host cells and vaccine  
 CC compositions comprising GapC polynucleotides or proteins. The vaccines  
 CC are used to treat or prevent a bacterial infection, especially a  
 CC streptococcal infection, and mastitis in particular (claimed). GapC  
 CC proteins are also used in claimed methods for detecting GapC antibodies,  
 CC and to raise antibodies that are used in claimed methods for detecting  
 CC GapC proteins. *S. uberis* is a common pathogen associated with mastitis in  
 CC cattle, horse, sheep and goat  
 XX  
 SQ Sequence 336 AA;

Query Match 68.5%; Score 1559.5; DB 5; Length 336;  
 Best Local Similarity 73.9%; Pred. No. 5.5e-121;  
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGLAFERRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTEV 87  
 1 MVVKVINGFGRIGLAFERRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTEV 60

QY 88 KEGGFEVNGNFIKVS 147  
 61 KDGGEVNGNFIKVS 81

QY 148 KDGPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANTDWATDGVEIV 207  
 82 -----NIDWATDGVEIV 95

QY 208 ATSFIAKKEAEKHLHANGAKKKVITAPGGNDVKTVVNTNHDILDGTETVVISGASCTTN 267  
 96 ATGFFAKKAAAEKHLHANGAKKKVITAPGGDDVKTVVNTNHDILDGTETVVISGASCTTN 155

DB  
 QY 268 CLAPMAKALHDAGIQLQGLMTTIHAYTQDMILDGPHRGGDLRRARAAGAANTIPVN 327  
 156 CLAPMAKALQDNGFVQQLGLMTTIHAYTQDMILDGPHRGGDLRRARAAGAANTIPVN 215

QY 328 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASND 387  
 216 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVAVLEKETSVEINAAMKAASND 275

QY 388 DPTVSSDIVGVSYGSLSFDATOTIKVMEVDGSQLVKVVSWYDNEMSYTAQLV 447  
 276 DPTVSSDIVGMAYSLSLFDATQTKVQTVDGNQLVKVVSWYDNEMSYTAQLV 335

QY 448 K 448  
 DB 336 K 336

RESULT 12  
 ID AAM50666  
 XX  
 AC AAM50666;  
 XX  
 DT 08-APR-2002 (first entry)  
 XX  
 DE Streptococcus agalactiae gapC plasmin binding protein AgalGapC.  
 XX  
 KW AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;  
 KW immunisation; mastitis; therapy.  
 XX  
 OS Streptococcus agalactiae.  
 XX  
 PN WO200196379-A2.  
 XX  
 PR 20-DEC-2001.  
 XX  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 XX  
 PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;  
 XX  
 DR 2002-130725/17.  
 XX  
 DR N-PSDB; ABA91250.  
 XX  
 PS Claim 1(c); Fig 3A-B; 107pp; English.  
 XX  
 PT Novel isolated GapC protein of *Streptococcus dysgalactiae*, *S. agalactiae*,  
 PT *S. uberis*, *S. paruberis*, or *S. iniae*, useful as vaccine component for  
 PT treating streptococcal infection which causes mastitis in vertebrates.  
 XX  
 PS The present sequence is that of the GapC plasmin binding protein  
 PT (*ÜberGapC*) of *Streptococcus uberis* ATCC 9927, encoded by the GapC gene  
 PT given in ABA91250. GapC, which has no signal sequence or membrane anchor  
 PT domain, is capable of eliciting an immune response in a vertebrate. The  
 PT invention provides the GapC genes and proteins of 5 *Streptococcus*  
 CC species, as well as recombinant vectors, host cells and vaccine  
 CC compositions comprising GapC polynucleotides or proteins. The vaccines  
 CC are used to treat or prevent a bacterial infection, especially a  
 CC streptococcal infection, and mastitis in particular (claimed). GapC  
 CC proteins are also used in claimed methods for detecting GapC antibodies,  
 CC and to raise antibodies that are used in claimed methods for detecting  
 CC GapC proteins. *S. uberis* is a common pathogen associated with mastitis in  
 CC cattle, horse, sheep and goat  
 XX  
 SQ Sequence 336 AA;

Query Match 68.5%; Score 1559.5; DB 5; Length 336;  
 Best Local Similarity 73.9%; Pred. No. 5.5e-121;  
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGLAFERRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTEV 87  
 1 MVVKVINGFGRIGLAFERRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTEV 60

QY 88 KEGGFEVNGNFIKVS 147  
 61 KDGGEVNGNFIKVS 81

QY 148 KDGPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANTDWATDGVEIV 207  
 82 -----NIDWATDGVEIV 95

QY 208 ATSFIAKKEAEKHLHANGAKKKVITAPGGNDVKTVVNTNHDILDGTETVVISGASCTTN 267  
 96 ATGFFAKKAAAEKHLHANGAKKKVITAPGGDDVKTVVNTNHDILDGTETVVISGASCTTN 155

DB  
 QY 268 CLAPMAKALHDAGIQLQGLMTTIHAYTQDMILDGPHRGGDLRRARAAGAANTIPVN 327  
 156 CLAPMAKALQDNGFVQQLGLMTTIHAYTQDMILDGPHRGGDLRRARAAGAANTIPVN 215

QY 328 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASND 387  
 216 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVAVLEKETSVEINAAMKAASND 275

QY 388 DPTVSSDIVGVSYGSLSFDATOTIKVMEVDGSQLVKVVSWYDNEMSYTAQLV 447  
 276 DPTVSSDIVGMAYSLSLFDATQTKVQTVDGNQLVKVVSWYDNEMSYTAQLV 335

QY 448 K 448  
 DB 336 K 336

RESULT 12  
 ID AAM50666  
 XX  
 AC AAM50666;  
 XX  
 DT 08-APR-2002 (first entry)  
 XX  
 DE Streptococcus agalactiae gapC plasmin binding protein AgalGapC.  
 XX  
 KW AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;  
 KW immunisation; mastitis; therapy.  
 XX  
 OS Streptococcus agalactiae.  
 XX  
 PN WO200196379-A2.  
 XX  
 PR 20-DEC-2001.  
 XX  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 XX  
 PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;  
 XX  
 DR 2002-130725/17.  
 XX  
 DR N-PSDB; ABA91250.  
 XX  
 PS Claim 1(c); Fig 3A-B; 107pp; English.  
 XX  
 PT Novel isolated GapC protein of *Streptococcus dysgalactiae*, *S. agalactiae*,  
 PT *S. uberis*, *S. paruberis*, or *S. iniae*, useful as vaccine component for  
 PT treating streptococcal infection which causes mastitis in vertebrates.  
 XX  
 PS The present sequence is that of the GapC plasmin binding protein  
 PT (*ÜberGapC*) of *Streptococcus uberis* ATCC 9927, encoded by the GapC gene  
 PT given in ABA91250. GapC, which has no signal sequence or membrane anchor  
 PT domain, is capable of eliciting an immune response in a vertebrate. The  
 PT invention provides the GapC genes and proteins of 5 *Streptococcus*  
 CC species, as well as recombinant vectors, host cells and vaccine  
 CC compositions comprising GapC polynucleotides or proteins. The vaccines  
 CC are used to treat or prevent a bacterial infection, especially a  
 CC streptococcal infection, and mastitis in particular (claimed). GapC  
 CC proteins are also used in claimed methods for detecting GapC antibodies,  
 CC and to raise antibodies that are used in claimed methods for detecting  
 CC GapC proteins. *S. uberis* is a common pathogen associated with mastitis in  
 CC cattle, horse, sheep and goat  
 XX  
 SQ Sequence 336 AA;

XX WPI; 2002-098051/13.  
 DR N-PSDB; ABA91329.  
 XX Novel GapC multiple epitope fusion polypeptide comprising antigenic  
 PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.  
 parauberis*, or *S. iniae* GapC protein useful for treating mastitis in  
 PT vertebrates.  
 XX Claim 8; Fig 2A-B; 116pp; English.  
 XX The present sequence is that of the GapC plasmin binding protein,  
 CC AgalGapC, of *Streptococcus agalactiae* ATCC 27541. The invention relates  
 CC to novel GapC multiple epitope fusion proteins that comprise epitopes  
 CC from 1 or more of *Streptococcus dysgalactiae*, *Streptococcus agalactiae*,  
 CC *Streptococcus parauberis*, *Streptococcus uberis* and *Streptococcus iniae*  
 CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).  
 CC Expression vectors and host cells for production of the multiple epitope  
 CC fusion proteins are provided. The multiple epitope proteins are used in  
 CC a vertebrate, especially a streptococcal infection, and particularly  
 CC in mastitis. They are also used in claimed methods of detecting  
 CC streptococcal antibodies. The multiple epitope proteins are capable of  
 CC eliciting broad immunity against a variety of streptococcal infections  
 CC while minimising the number of antigens present in the final formulation  
 CC and concomitantly reducing production costs  
 XX Sequence 336 AA;  
 Query Match 68.4%; Score 1557.5; DB 5; Length 336;  
 Best Local Similarity 73.9%; Pred. No. 8.1e-121;  
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;  
 Qy 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMMLAHLILKYDTTQGRFDGTVEV 87  
 Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMMLAHLILKYDTTQGRF----- 54  
 Qy 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVETVLEGTVKDGFFDVNGKFIKVSAP 147  
 Db 55 ----- 54  
 Qy 148 KDPEQIDWATDGVETVLEGTVKDGFFEVNGQFVKAEREPANIDWATDGVETVLEGTVKDGFFDVNGKFIKVSAP 207  
 Db 55 ----- 54  
 Qy 208 ATSFEEAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFVNTNHDILDGTETVSGASCTTN 267  
 Db 96 ATGFPEASKEKAGQHIIHENGAKKVVITAPGGNDVKTVFVNTNHDILDGTETVSGASCTTN 155  
 Qy 268 CLAPMAKALHDAGFGIQKGGLMTTIHAYTGDCMILDGPHRGDDLRRARAAGAAANIVPNSTGAA 327  
 Db 156 CLAPMAKALQDNFGVKQGLMTTIHAYTGDCMILDGPHRGDDLRRARAAGAAANIVPNSTGAA 215  
 Qy 328 KAIGLVIPELNGKLDGAAQQRVPVPTGSVTELYVTLDKNVSVDEINAAAMKAASNDSFEGTYE 387  
 Db 216 KAIGLVIPELNGKLDGAAQQRVPVPTGSVTELVATELKDVTEVNAAMKAANDSYGYTE 275  
 Qy 388 DPIVSSDIVGVSYGSLEFDATQTKVMEVGDGSQLVKVVSWYDNEMSYTAQLVRLTEYFAKIA 447  
 Db 276 DPIVSSDIVGSIYGSLEFDATQTKVQTVDGNQLVKVVSWYDNEMSYTSVQLRTEYFAKIA 335  
 Qy 448 K 448  
 Db 336 K 336  
 Qy RESULT 13  
 AAM50640  
 ID AAM50640 standard; protein; 336 AA.  
 XX AC AAM50640;  
 DT 04-APR-2002 (first entry)

XX Streptococcus agalactiae gapC plasmin binding protein.  
 DE GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;  
 XX KW diagnosis; therapy.  
 KW OS Streptococcus agalactiae.  
 PT PN WO200196381-A2.  
 XX PD 20-DEC-2001.  
 XX PF 11-JUN-2001; 2001WO-CA0000838.  
 XX PR 12-JUN-2000; 2000US-0211022P.  
 XX PA (UYSA-) UNIV SASKATCHEWAN.  
 XX PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;  
 XX DR WPI; 2002-130725/17.  
 XX DR N-PSDB; ABA91249.  
 XX PS Claim 1 (b); Fig 2A-B; 107pp; English.  
 XX CC The present sequence is that of the GapC plasmin binding protein  
 CC (AgalGapC) of *Streptococcus agalactiae* ATCC 27541, encoded by the GapC  
 CC gene given in ABA91249. GapC, which has no signal sequence or membrane  
 CC anchor domain, is capable of eliciting an immune response in a  
 CC vertebrate. The invention provides the GapC genes and proteins of 5  
 CC *Streptococcus* species, as well as recombinant vectors, host cells and  
 CC vaccine compositions comprising GapC polynucleotides or proteins. The  
 CC vaccines are used to treat or prevent a bacterial infection, especially a  
 CC streptococcal infection, and mastitis in particular (claimed). GapC  
 CC proteins are also used in claimed methods for detecting GapC antibodies,  
 CC and to raise antibodies that are used in claimed methods for detecting  
 CC GapC proteins. *S. agalactiae* is a common pathogen associated with  
 CC mastitis in cattle, horse, sheep and goat. It also causes septicemia,  
 CC meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,  
 CC abscesses, spontaneous abortion, etc  
 XX SQ Sequence 336 AA;  
 Query Match 68.4%; Score 1557.5; DB 5; Length 336;  
 Best Local Similarity 73.9%; Pred. No. 8.1e-121;  
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;  
 Qy 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMMLAHLILKYDTTQGRFDGTVEV 87  
 Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMMLAHLILKYDTTQGRF----- 54  
 Qy 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVETVLEGTVKDGFFDVNGKFIKVSAP 147  
 Db 55 ----- 54  
 Qy 148 KDPEQIDWATDGVETVLEGTVKDGFFEVNGQFVKAEREPANIDWATDGVETVLEGTVKDGFFDVNGKFIKVSAP 207  
 Db 55 ----- 54  
 Qy 208 ATSFEEAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFVNTNHDILDGTETVSGASCTTN 267  
 Db 96 ATGFPEASKEKAGQHIIHENGAKKVVITAPGGNDVKTVFVNTNHDILDGTETVSGASCTTN 155  
 Qy 268 CLAPMAKALHDAGFGIQKGGLMTTIHAYTGDCMILDGPHRGDDLRRARAAGAAANIVPNSTGAA 327  
 Db 156 CLAPMAKALQDNFGVKQGLMTTIHAYTGDCMILDGPHRGDDLRRARAAGAAANIVPNSTGAA 215  
 Qy 328 KAIGLVIPELNGKLDGAQQRVPVPTGSVTELYVTLDKNVSVDEINAAAMKAASNDSFEGTYE 387  
 Db 216 KAIGLVIPELNGKLDGAQQRVPVPTGSVTELVATELKDVTEVNAAMKAANDSYGYTE 275  
 Qy 388 DPIVSSDIVGVSYGSLEFDATQTKVMEVGDGSQLVKVVSWYDNEMSYTAQLVRLTEYFAKIA 447  
 Db 276 DPIVSSDIVGSIYGSLEFDATQTKVQTVDGNQLVKVVSWYDNEMSYTSVQLRTEYFAKIA 335  
 Qy 448 K 448  
 Db 336 K 336  
 Qy RESULT 13  
 AAM50640  
 ID AAM50640 standard; protein; 336 AA.  
 XX AC AAM50640;  
 DT 04-APR-2002 (first entry)

XX Streptococcus agalactiae gapC plasmin binding protein.  
 DE GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;  
 XX KW diagnosis; therapy.  
 KW OS Streptococcus agalactiae.  
 PT PN WO200196381-A2.  
 XX PD 20-DEC-2001.  
 XX PF 11-JUN-2001; 2001WO-CA0000838.  
 XX PR 12-JUN-2000; 2000US-0211022P.  
 XX PA (UYSA-) UNIV SASKATCHEWAN.  
 XX PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;  
 XX DR WPI; 2002-130725/17.  
 XX DR N-PSDB; ABA91249.  
 XX PS Claim 1 (b); Fig 2A-B; 107pp; English.  
 XX CC The present sequence is that of the GapC plasmin binding protein  
 CC (AgalGapC) of *Streptococcus agalactiae* ATCC 27541, encoded by the GapC  
 CC gene given in ABA91249. GapC, which has no signal sequence or membrane  
 CC anchor domain, is capable of eliciting an immune response in a  
 CC vertebrate. The invention provides the GapC genes and proteins of 5  
 CC *Streptococcus* species, as well as recombinant vectors, host cells and  
 CC vaccine compositions comprising GapC polynucleotides or proteins. The  
 CC vaccines are used to treat or prevent a bacterial infection, especially a  
 CC streptococcal infection, and mastitis in particular (claimed). GapC  
 CC proteins are also used in claimed methods for detecting GapC antibodies,  
 CC and to raise antibodies that are used in claimed methods for detecting  
 CC GapC proteins. *S. agalactiae* is a common pathogen associated with  
 CC mastitis in cattle, horse, sheep and goat. It also causes septicemia,  
 CC meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,  
 CC abscesses, spontaneous abortion, etc  
 XX SQ Sequence 336 AA;  
 Query Match 68.4%; Score 1557.5; DB 5; Length 336;  
 Best Local Similarity 73.9%; Pred. No. 8.1e-121;  
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;  
 Qy 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMMLAHLILKYDTTQGRFDGTVEV 87  
 Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMMLAHLILKYDTTQGRF----- 54  
 Qy 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVETVLEGTVKDGFFDVNGKFIKVSAP 147  
 Db 55 ----- 54  
 Qy 148 KDPEQIDWATDGVETVLEGTVKDGFFEVNGQFVKAEREPANIDWATDGVETVLEGTVKDGFFDVNGKFIKVSAP 207  
 Db 55 ----- 54  
 Qy 208 ATSFEEAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFVNTNHDILDGTETVSGASCTTN 267  
 Db 96 ATGFPEASKEKAGQHIIHENGAKKVVITAPGGNDVKTVFVNTNHDILDGTETVSGASCTTN 155  
 Qy 268 CLAPMAKALHDAGFGIQKGGLMTTIHAYTGDCMILDGPHRGDDLRRARAAGAAANIVPNSTGAA 327  
 Db 156 CLAPMAKALQDNFGVKQGLMTTIHAYTGDCMILDGPHRGDDLRRARAAGAAANIVPNSTGAA 215  
 Qy 328 KAIGLVIPELNGKLDGAQQRVPVPTGSVTELYVTLDKNVSVDEINAAAMKAASNDSFEGTYE 387  
 Db 216 KAIGLVIPELNGKLDGAQQRVPVPTGSVTELVATELKDVTEVNAAMKAANDSYGYTE 275  
 Qy 388 DPIVSSDIVGVSYGSLEFDATQTKVMEVGDGSQLVKVVSWYDNEMSYTAQLVRLTEYFAKIA 447  
 Db 276 DPIVSSDIVGSIYGSLEFDATQTKVQTVDGNQLVKVVSWYDNEMSYTSVQLRTEYFAKIA 335  
 Qy 448 K 448  
 Db 336 K 336  
 Qy RESULT 13  
 AAM50640  
 ID AAM50640 standard; protein; 336 AA.  
 XX AC AAM50640;  
 DT 04-APR-2002 (first entry)

Db 216 KAIGLVIPELNGKLIDGAAQRVPVPTGSVTELVATLEKDVTVEEVNAAMKAANDSYGYTE 275  
 QY 388 DPIVSSDIVGVSYGSILFDATQTKVMEVIDGSQLVKVVSYDNEMSYTAQLVRLTEYFAKIA 447  
 Db 276 DPIVSSDIVGVSYGSILFDATQTKVOTVDGNLQVKVVSYDNEMSYTAQLVRLTEYFAKIA 335  
 QY 448 K 448  
 Db 336 K 336

RESULT 14  
 ABU02516 standard; protein; 335 AA.  
 XX  
 AC ABU02516;  
 XX DT 23-OCT-2003 (revised)  
 DT 11-FEB-2003 (first entry)  
 XX  
 S. pneumoniae type 4 strain protein from coding region #2094.  
 XX  
 DE Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
 KW gene therapy; vaccine.  
 XX  
 OS Streptococcus pneumoniae; type 4 strain.  
 XX  
 PN WO200277021-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-IB0002163.  
 XX  
 PR 27-MAR-2001; 2001GB-00007658.  
 XX  
 PA (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Masignani V, Tettelin H, Fraser C;  
 XX  
 DR WPI; 2003-040579/03.  
 DR N-PSDB; ABX07806.  
 XX  
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
 PT useful as medicaments for treating or preventing a disease or infection  
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or  
 PT ear infection.  
 XX  
 PS Claim 1; SEQ ID NO 4188; 56pp; English.

CC The invention relates to a protein comprising or having at least 50%  
 CC identity to any of the 2469 amino acid sequences, identified in the  
 CC specification (available on a computer readable format), or its fragment,  
 CC expressed from 2469 of 2489 identified DNA coding regions from the  
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as  
 CC ABS56454. Also included are an antibody which binds one of the proteins,  
 CC treating a patient by administering the protein, DNA or antibody (in a  
 CC composition), a kit comprising first and second primers, which are the  
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a  
 CC sequence not defined in the specification, for amplifying a target  
 CC sequence contained within a Streptococcus nucleic acid sequence, where  
 CC the first primer is substantially complementary to the target sequence  
 CC and the second primer is substantially complementary to the complement of  
 CC the target sequence, and where the parts of the primers having  
 CC substantial complementarity define the termini of the target sequence to  
 CC be amplified, assay comprising contacting a test compound with the  
 CC protein, and determining whether the test compound binds to the protein  
 CC and a Streptococcus pneumoniae bacterium, where one or more genes  
 CC encoding the proteins has been rendered inactive. The proteins, nucleic  
 CC acid molecules, antibody and compositions are useful as medicaments for  
 CC treating or preventing a disease or infection due to streptococcus  
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis

CC media or ear infection. They are also useful in developing vaccines,  
 CC diagnostics and antibiotics. The methods are useful for identifying  
 CC immunodominant proteins. The present sequence is one of the 2469 proteins  
 CC expressed by the identified coding regions from the genomic sequence.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences). (Updated on 23-OCT-2003 to  
 CC standardise OS field)

XX  
 SQ Sequence 335 AA;

Query	Match	Score	DB	Length
QY	28 MVVKVGINGFGRIGRLAFRRIONQVEVTRINDLTPNMLAHLLKYDTTQGRFDGTVEV	67.4%	6	335;
Db	1 MVVRVGINGFGRIGRLAFRRIONQVEVTRINDLTPVMLAHLLKYDTTQGRFDGTVEV	73.6%	6	335;
QY	88 KEGGFEVNGNFIKVSAERDPENIDWATDGWEIVLEALEGTVEVKDGGFDVNGKPIKVSAE	147	61	335;
Db	82 KDPEQIDWATDGWEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVEIVLE	207	82	335;
QY	208 ATSEFAKKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVISGASCTTN	267	96	335;
Db	96 ATGFFFAKKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVISGASCTTN	154	96	335;
QY	268 CLAPMAKALHDAGFIQKGLMTTIIHAYTGQMLDGPGRGGDLRRARAAGAANIVPNSTGAA	327	155	335;
Db	155 CLAPMAKALQDNFNGVVEGLMTTIIHAYTGQMLDGPGRGGDLRRARAAGAANIVPNSTGAA	214	155	335;
QY	328 KAIGLVIPELINGKLQDGAAQRPVPPTGSVTELVVTLDKNVSVDENIAAMKAASNDSFEGYTE	387	215	335;
Db	215 KAIGLVIPELINGKLQDGSAAQRVPPTPTGSVTELVAVLEKNTVDEVNAAMKAASNESYGYTE	274	215	335;
QY	388 DPIVSSDIVGVSYGSILFDATQTKYMEVDGSQQLVKKVSYWDNEMSYTAQLVRLTEYFAKIA	447	448	335;
Db	448 K 448	448	448	335;

RESULT 15  
 AAU37576  
 ID AAU37576 standard; protein; 359 AA.  
 XX  
 AC AAU37576;  
 XX DT 14-FEB-2002 (first entry)  
 XX DE Streptococcus pneumoniae cellular proliferation protein #5.  
 XX OS Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 XX KW antibacterial; drug design.  
 XX OS Streptococcus pneumoniae.  
 XX PN WO200170955-A2.  
 XX PD 27-SEP-2001.  
 XX PF 21-MAR-2001; 2001WO-US009180.  
 XX PR 21-MAR-2000; 2000US-0191078P.  
 XX PR 23-MAY-2000; 2000US-0206848P.  
 XX PR 26-MAY-2000; 2000US-022772P.  
 XX PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI; 2001-611495/70.  
 XX  
 DR N-PSDB; AAS55435.

PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX PS Example 3; SEQ ID NO 13169; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

XX SQ Sequence 359 AA;

Query	Match	Score	DB	Length
QY	28 MVVKVGINGFGRIGRLAARRRIONVEVTRINDLTPNMLAHLILKYDTTQGRFDGTVEV	67.4%	4	359;
Db	25 MVVKVGINGFGRIGRLAARRRIONVEVTRINDLTPVMLAHLILKYDTTQGRFDGTVEV	73.6%	119;	
QY	88 KEGGFEVNGGNFIKVSaERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSaE	87	2;	
Db	85 KEGGFEVNGKFIKVSaERDP-	84		
QY	148 KDPEQIDWATDGVEIVLEIDGTVEVKEGGEVNGQFVKVSAEREPANIDWATDGVEIVLE	147	105	
Db	106 -	-	-	-
QY	208 ATSFRAKKEAAEKLH-HANGAKKKVITAPEGNDVKTFFNTNHDILDGTETVISGASCTN	207	119	
Db	120 ATGFFAKKEAAEKLH-KGGAKKKVITAPEGNDVKTFFNTNHDILDGTETVISGASCTN	267		
QY	268 CLAPMAKALHDAGFIQKGGLMTTHAYTGDQMLILDGPHRGDRLRARAAGAANIVPNSTGAA	178	327	
Db	179 CLAPMAKALQDNFGVVEGLMTTHAYTGDQMLILDGPHRGDRLRARAAGAANIVPNSTGAA	238		
QY	328 KAIGLVIPELNGKLQDGAAQRVPPTGSVTELVYVTLDKNVSVDEVAAMKAASNDSFQYTE	327	387	
Db	239 KAIGLVIPELNGKLQDGSAQRVPPTGSVTELVAVLEKNTVDEVAAMKAASNESYGYTE	358		
QY	388 DPIVSSDIVGVSYGSLFDATQTKVMEVGDGSQLVKVVSWYDNEMSYTAQLVRLTEYFAKIA	358	447	
Db	299 DPIVSSDIVGMSYGSLFDATQTKVLDVGDGKQLVKVVSWYDNEMSYTAQLVRLTEYFAKIA	448		
QY	448 K 448	359	K 359	
Db				

This Page Blank (uspto)

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:00:07 ; Search time 34 Seconds  
 (without alignments)  
 680.248 Million cell updates/sec

Title: US-10-650-369-22  
 Perfect score: 2278  
 Sequence: 1 MKKITGIIILLLAVILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing First 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/iaa/PC7US\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2278	100.0	448	4	US-09-878-766A-22	Sequence 22, Appl
2	1656.5	72.7	336	4	US-09-878-766A-12	Sequence 12, Appl
3	1652.5	72.5	336	1	US-07-928-462-2	Sequence 2, Appl
4	1652.5	72.5	336	3	US-08-273-247-2	Sequence 2, Appl
5	1559.5	68.5	336	4	US-09-878-766A-16	Sequence 16, Appl
6	1557.5	68.4	336	4	US-09-878-766A-14	Sequence 14, Appl
7	1534.5	67.4	336	4	US-09-878-766A-20	Sequence 20, Appl
8	1530.5	67.2	336	4	US-09-878-766A-18	Sequence 18, Appl
9	1491	65.5	333	3	US-08-961-083-54	Sequence 54, Appl
10	1491	65.5	333	4	US-09-536-784-54	Sequence 54, Appl
11	1282	56.3	346	4	US-09-134-000C-4400	Sequence 4400, Appl
12	1270	55.8	333	4	US-09-107-532A-4769	Sequence 4769, Appl
13	1138.5	50.0	348	4	US-09-134-001C-5513	Sequence 5513, Appl
14	923.5	40.5	357	4	US-09-134-000C-4229	Sequence 4229, Appl
15	918.5	40.3	340	4	US-09-634-238-233	Sequence 233, Appl
16	905.5	39.7	349	4	US-09-489-039A-13939	Sequence 13939, Appl
17	84.9	37.3	335	6	5290690-9	Patent No. 5290690
18	83.8	36.8	334	6	5290690-10	Patent No. 5290690
19	80.3	35.3	340	3	US-09-095-855-205	Sequence 205, Appl
20	80.3	35.3	340	4	US-09-205-426-205	Sequence 205, Appl
21	78.0	34.2	334	6	5290690-11	Patent No. 5290690
22	776.5	34.1	335	4	US-09-489-039A-7679	Sequence 7679, Appl
23	758.5	33.3	282	4	US-09-107-532A-4057	Sequence 4057, Appl
24	747.5	32.8	340	4	US-09-543-681A-9083	Sequence 8083, Appl
25	719	31.6	352	4	US-09-489-039A-9410	Sequence 9410, Appl
26	690	30.3	336	4	US-09-198-452A-662	Sequence 662, Appl
27	683.5	30.0	338	4	US-09-091-725-10	Sequence 10, Appl

#### ALIGNMENTS

RESULT 1 US-09-878-766A-22						
; Sequence 22, Application	US/09878766A					
; Patent No. 6660270						
; GENERAL INFORMATION:						
; APPLICANT: Perez-Casal, Jose						
; APPLICANT: Fontaine, Michael						
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN						
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION						
; FILE REFERENCE: 9000-0057						
; CURRENT APPLICATION NUMBER: US/09/878, 766A						
; CURRENT FILING DATE: 2001-09-10						
; NUMBER OF SEQ ID NOS: 22						
; SOFTWARE: PatentIn Ver. 2.0						
; SEQ ID NO 22						
; LENGTH: 448						
; TYPE: PRT						
; ORGANISM: Artificial Sequence						
; FEATURE:						
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple						
; OTHER INFORMATION: epitope fusion protein						
; US-09-878-766A-22						

Query	Match	Score 100.0%	DB 4
Best Local Similarity	100.0%	Pred. No. 1.7e-206;	
Matches 448; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKKITGILLLAVILSACQANYGSGMVVKGINGFGRIGLAFRRIONVEGVETRIN 60	
Db	1	MKKITGILLLAVILSACQANYGSGMVVKGINGFGRIGLAFRRIONVEGVETRIN 60	
QY	61	DLTDPNMLLAHLKYDTTQGRFDGTVKEGGFVNNGNFIVKSAERDPENIDWATDGVETIV 120	
Db	61	DLTDPNMLLAHLKYDTTQGRFDGTVKEGGFVNNGNFIVKSAERDPENIDWATDGVETIV 120	
QY	61	LEALEGTVEVKDGFDVNGKFLIVKSAEKDPEQIDWATDGVETIVKEGGFVN 180	
Db	61	LEALEGTVEVKDGFDVNGKFLIVKSAEKDPEQIDWATDGVETIVKEGGFVN 180	
QY	181	GQFVKVSAEREPAKKEAAEKHLHANGAKKKVVTAPGGNDV 240	
Db	181	KTVVVENTNHDILDGTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIAHTGQDML 300	
QY	241	KTVVVENTNHDILDGTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTIAHTGQDML 300	
Db	241	DGPHRGDDLRARRAGANIVPNSTGAAKAIGLVIPEINGKLDGAACRQVPVPTGSVTELVY 360	
QY	301		

Db 301 DGPFRGGDLRRARRAGAAANIVPNSTGAAKAIGLVIPELNGKLQAAQRVPEVPTGSVTELVV 3.60

QY 361 TLDKNVSDEINAAKAASNDSFGYTEDPIVSSDILGVSYGSLFDATQTKYMEVDGSQLV 4.20

Db 361 TLDKNVSDEINAAKAASNDSFGYTEDPIVSSDILGVSYGSLFDATQTKYMEVDGSQLV 4.20

QY 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 4.48

Db 421 KVVSWYDNEMSYTAQLVRTLEYFAKIAK 4.48

RESULT 2  
US-09-878-766A-12  
; Sequence 12, Application US/09878766A

; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus dysgalactiae  
; US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 4; Length 336;  
Best Local Similarity 79.6%; Pred. No. 5e-148; Mismatches 0; Indels 1; Gaps 1;  
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVGINGFGRIGRLAFRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
1 MVVKVGINGFGRIGRLAFRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFFEVNGNFIKVSAERDPEIDWATDGVETVLEALEGTVEVRDGGFDVNGKFKIKVSAE 147  
61 KEGGFFEVNGNFIKVSAERDPE-----81

QY 148 KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVETVLE 207  
82 -----81

Db 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 267  
96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 155

Db 268 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 327  
156 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 215

Db 328 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 387  
216 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 275

QY 388 DPIVSSDILGVSYGSLFDATQTKYMEVDGSQLVKVKVSYWDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDILGVSYGSLFDATQTKYMEVDGSQLVKVKVSYWDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448  
Db 336 K 336

Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 155

Db 148 KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVETVLE 207  
82 -----81

Db 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 267  
156 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 387  
Db 216 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 275

GENERAL INFORMATION:  
; APPLICANT: Boyle, Michael D.P.  
; APPLICANT: Lottenberg, Richard  
; APPLICANT: Broder, Christopher C.  
; APPLICANT: von Mering, Gregory O.  
; TITLE OF INVENTION: Bacterial Plasmin Receptors as  
; TITLE OF INVENTION: Fibrinolytic Agents  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/928,462  
; FILING DATE: 19920810  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/524,411  
; FILING DATE: 16-MAY-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/330,849  
; FILING DATE: 29-MAR-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-928-462-2

Query Match 72.5%; Score 1652.5; DB 1; Length 336;  
Best Local Similarity 79.1%; Pred. No. 1.2e-147; Mismatches 2; Indels 85; Gaps 1;  
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVGINGFGRIGRLAFRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87  
1 MVVKVGINGFGRIGRLAFRRIQNVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFFEVNGNFIKVSAERDPEIDWATDGVETVLEALEGTVEVRDGGFDVNGKFKIKVSAE 147  
61 KEGGFFEVNGNFIKVSAERDPE-----81

QY 148 KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVETVLE 207  
82 -----81

Db 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 267  
96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 327  
156 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 215

Db 328 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 387  
Db 216 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 275

QY 388 DPIVSSDILGVSYGSLFDATQTKYMEVDGSQLVKVKVSYWDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDILGVSYGSLFDATQTKYMEVDGSQLVKVKVSYWDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448  
Db 336 K 336

Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 155

Db 148 KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVETVLE 207  
82 -----81

Db 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 267  
156 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 387  
Db 216 KAIGLVIPELNGKLQAAQRVPVPTGSVTELVTLDRNVSVDINAAMKAASNDSEGYTE 275

RESULT 3  
US-07-928-462-2  
; Sequence 2, Application US/07928462  
; Patent No. 5328996

QY 388 DPIVSSDIVGVSYGSILFDATQTKVMEVDGSSQLVKVVSWNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGVSYGSILFDATQTKVMEVDGSSQLVKVVSWNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448  
Db 336 K 336

RESULT 4  
US-08-273-247-2  
; Sequence 2, Application US/08273247  
; GENERAL INFORMATION:  
; APPLICANT: Boyle, Michael D.P.  
; APPLICANT: Lottenberg, Richard  
; APPLICANT: Broder, Christopher C.  
; APPLICANT: von Mering, Gregory O.  
; TITLE OF INVENTION: Bacterial Plasmin Receptors as  
; TITLE OF INVENTION: Fibrinolytic Agents  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/273,247  
; FILING DATE: 29-MAR-1989  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/524,411  
; FILING DATE: 16-MAY-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/330,849  
; FILING DATE: 29-MAR-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-273-247-2

QY 208 ATSFIAKKEAEKHLHANGAKKVVITAPGGNDVKVVNTNHDILDGTETVVISGASCTN 267  
Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGGNDVKVVNTNHDILDGTETVVISGASCTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGHRRAGAANIVPNSTGAA 327  
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGHRRAGAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDAQAQRVPVPTGSVTELVTLDKNVSDEINAAMKAASNDSGFYTE 387  
Db 216 KAIGLVIPELNGKLDAQAQRVPVPTGSVTELVTLDKNVSDEINAAMKAASNDSGFYTE 275

QY 388 DPIVSSDIVGVSYGSILFDATQTKVMEVDGSSQLVKVVSWNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGVSYGSILFDATQTKVMEVDGSSQLVKVVSWNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448  
Db 336 K 336

RESULT 5  
US-09-878-766A-16  
; Sequence 16, Application US/09878766A  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
; FILE REFERENCE: 9000-0057  
; CURRENT APPLICATION NUMBER: US/09/878,766A  
; CURRENT FILING DATE: 2001-09-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus uberis  
; US-09-878-766A-16

Query Match 68.5%; Score 1559.5; DB 4; Length 336;  
Best Local Similarity 73.9%; Pred. No. 7.1e-139; Mismatches 15; Indels 85; Gaps 1;  
Matches 311; Conservative

QY 28 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIFKVSAAE 147  
Db 61 KDGFFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIFKVSAAE 81

QY 148 KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGVEIVLE 207  
Db 82 NIDWATDGVEIVLE 95

QY 208 ATSFIAKKEAEKHLHANGAKKVVITAPGGNDVKVVNTNHDILDGTETVVISGASCTN 267  
Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGGDDVKTVVNTNHDILDGTETVVISGASCTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGHRRAGAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFGVKQGLMTTIHAYTGDQMLDGHRRAGAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDAQAQRVPVPTGSVTELVTLDKNVSDEINAAMKAASNDSGFYTE 387  
Db 216 KAIGLVIPELNGKLDAQAQRVPVPTGSVTELVALEKETSVEINAAMKAASNDSGFYTE 275

QY 388 DPIVSSDIVGVSYGSILFDATQTKVMEVDGSSQLVKVVSWNEMSYTAQLVRTLEYFAKIA 447  
Db 82 NIDWATDGVEIVLE 95

Db 276 DPIVSDIIGMAYGSLFDATQTQVQTVGDNQLVKVVSWYDNEMSYTAQLVRTLEYFAKIA 335  
 Qy 448 K 448  
 Db 336 K 336

RESULT 6  
 US-09-878-766A-14  
 ; Sequence 14, Application US/09878766A  
 ; Patent No. 6660270  
 ; GENERAL INFORMATION:  
 ;

; APPLICANT: Potter, Andrew A.

; APPLICANT: Perez-Casal, Jose

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN

; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057

; CURRENT APPLICATION NUMBER: US/09/878,766A

; CURRENT FILING DATE: 2001-09-10

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 14

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Streptococcus agalactiae

Query Match 68.4%; Score 1557.5; DB 4; Length 336;  
 Best Local Similarity 73.9%; Pred. No. 1.1e-138;  
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

Qy 28 MVVKVQINGFGRIGRLAFRRIONVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTVEV 87  
 Db 1 MVVKVQINGFGRIGRLAFRRIONVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTVEV 60

Qy 88 KEGGFEVNGNFIKVSAERDPENIDWATDGVEIVLEALEGTVEVKDGFEVNGKFIKVSAE 147  
 Db 61 KDGGEFVNGS-----70

Qy 148 KDPEQIDWATDGVEIVLEIDGTVEVKDGFEVNGQFVKSAAEREPANIDWATDGVEIVLE 207  
 Db 71 -----FVKVSAEREPANIDWATDGDIVLE 95

Qy 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTFFNTNEDILDGTETVISGASCTTN 267  
 Db 96 ATGFFAASKAAAEQHIIHANGAKKVVITAPGNDVKTFFNTNHDILDGTETVISGASCTTN 155

Qy 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHGRGGDLRRARAGAANTIPNSTGAA 327  
 Db 156 CLAPMAKALQDNFGVKQGLMTTHAYTGDQMLDGHGRGGDLRRARAAAANIPNSTGAA 215

Qy 328 KAIGLVIPELINGKLDGAQRVPPVPTGSVTELVVTLDKNVSVDEINAAMKAANSDFGYTE 387  
 Db 216 KAIGLVIPELINGKLDGAQRVPPVPTGSVTELVAVLEKDTSEEINAAMKAANDSGYTE 275

Qy 388 DPIVSSDIVGVSYGSLFDATQTQVMEVDGSQLVKVVSWYDNEMSYTAQLVRLTEYFAKIA 447

Db 276 DAIVSSDIVGISYGSLFDATQTQVTDGNQVLKVVVSWYDNEMSYTAQLVRLTEYFAKIA 335

Qy 448 K 448

Db 336 K 336

RESULT 8  
 US-09-878-766A-18  
 ; Sequence 18, Application US/09878766A  
 ; Patent No. 6660270  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Potter, Andrew A.  
 ; APPLICANT: Perez-Casal, Jose  
 ; APPLICANT: Fontaine, Michael  
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
 ; FILE REFERENCE: 9000-0057  
 ; CURRENT APPLICATION NUMBER: US/09/878,766A  
 ; CURRENT FILING DATE: 2001-09-10  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 18  
 ; LENGTH: 336

RESULT 7  
 US-09-878-766A-20  
 ; Sequence 20, Application US/09878766A  
 ; Patent No. 6660270  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Potter, Andrew A.  
 ; APPLICANT: Perez-Casal, Jose  
 ; APPLICANT: Fontaine, Michael

Query Match 67.2%; Score 1530.5; DB 4; Length 336;  
 Best Local Similarity 73.4%; Pred. No. 3.9e-136;  
 US-09-878-766A-18

Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2; STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-961-083-54

Query 28 MVVKVINGFGRIGRLAARRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87  
 Database 1 MVVKVINGFGRIGRLAARRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFD---- 55

Query 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALLEGTVEVKDGGFDVNGKFIKVSAAE 147  
 Database 56 -----GTVEVKDGGFDVNGKFIKVSAAE 77

Query 148 KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGVEIVLE 207  
 Database 78 KDPEQ-----IDWATDGVEIVLE 95

Query 208 ATSFPEAKKEAAEKHLHANGAKKUVITAPGGNDVKTUVFNTNHDILDGTETVSGASCTN 267  
 Database 96 ATGFFPEAKKEAAEKHLHENGAKKVITAPGGDDVKTUVFNTNHDILDGTETVSGASCTN 155

Query 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGQMLILDGPHRGGLRRARAGAAANTVPNSTGAA 327  
 Database 156 CLAPMAKALQDNFGVKQGLMTTIHAYTGQMLILDGPHRGGLRRARAGANNVPNSTGAA 215

Query 328 KAIGLVIPELNGKLDGAAQRVPPVPTGSVTELVVTLDRKNSVDEINAAMKAASNDSEGYTE 387  
 Database 216 KAIGLVIPELNGKLDGAAQRVPPVPTGSVTELVAVLNKETSVEEINSVMKAANDSEGYTE 275

Query 388 DPIVSSDIVGVSYGSLSFDATQTKVMEVDGSQQLVKVVSWDNEMSYTAQLVRLTEYFAKIA 447  
 Database 276 DPIVSSDIVGMSFGSLFDATQTKVQTVGDGNQLVKVVSWDNEMSYTAQLDRTLEYFAKIA 335

Query 448 K 448  
 Database 336 K 336

RESULT 9  
 US-08-961-083-54  
 Sequence 54, Application US/08961083  
 Patent No. 6159469  
 GENERAL INFORMATION:  
 APPLICANT: Choi et. al.  
 TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
 NUMBER OF SEQUENCES: 452  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/961,083  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brookes, A. Anders  
 REGISTRATION NUMBER: 36,373  
 REFERENCE/DOCKET NUMBER: PB340P2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 54:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 333 amino acids  
 TYPE: amino acid

RESULT 10  
 US-09-536-784-54  
 Sequence 54, Application US/09536784  
 Patent No. 6573082  
 GENERAL INFORMATION:  
 APPLICANT: Choi et. al.  
 TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
 NUMBER OF SEQUENCES: 452  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/536,784  
 FILING DATE: OCT 30-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Michelle S. Marks  
 REGISTRATION NUMBER: 41,971  
 REFERENCE/DOCKET NUMBER: PB340P3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 54:  
 INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 333 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 54:  
 US-09-536-784-54

Query Match 65.5%; Score 1491; DB 4; Length 333;  
 Best Local Similarity 72.4%; Pred. No. 2.1e-132;  
 Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;

Db 74 HEGSFNVNGKEIKVLANRNPPEELPW----- 98  
 QY 148 KDPEQIDWATDGVEIVLEIDGTEVKEGGFEVNGQFVVKVSAEREPANIDWATDGVEIVLE 207  
 Db 99 ----- 1.08  
 QY 208 ATSEFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVSGASCTTN 267  
 Db 109 CTGFFTSKEAAERHLTA-GAKRVRVISAPGGNDVPTIVNTNHELTGEETVSGASCTTN 167  
 QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAANIVPNSTGAA 327  
 Db 168 CLAPMAKALHDNFGVVEGLMTTIHAYTGDQMILDGPHPKGDERRARAAAANTVPNSTGAA 227  
 QY 328 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSYDEINAAMKAASNDSGFGYTE 387  
 Db 228 KAIGLVIPELNGKLDGAQRVPVATGSLTELVTLDEVNAVMKAANESYGYNT 287  
 QY 388 DPIVSSDIVGVSYGSLSFDATQTKVMEVDGSQQLVKVVSWDNEMSYTAQLVRLTLEYFAKI 446  
 Db 288 DEIVSSDIVGMYGSLSFDATQTKVMTVGDKQLVKTVWAWDNEMSYTAQLVRLTLEYFANL 346

Db 81 -----  
 -QIDWATDGVEIVLEIDGTEVKEGGFEVNGQFVVKVSAEREPANIDWATDGVEIVLEA 95  
 RESULT 12  
 US-09-107-532A-4769  
 ; Sequence 4769, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD/ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107-532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085, 598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Arinille, Pamela Deneke  
 ; REGISTRATION NUMBER: 40, 489  
 ; REFERENCE/DOCKET NUMBER: GTC-012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781) 893-5007  
 ; TELEFAX: (781) 893-8277  
 ; INFORMATION FOR SEQ ID NO: 4769:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 333 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Enterococcus faecium  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1...333  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4769:  
 ; US-09-107-532A-4769

Query Match 56.3%; Score 1282; DB 4; Length 346;  
 Best Local Similarity 62.5%; Pred. No. 1.2e-112;  
 Matches 262; Conservative 25; Mismatches 46; Indels 86; Gaps 3;

QY 28 MVVKVGINGFGRIGRLAFRRRIQNVGEGEVTRINDLTDPNMLAHLLKYDTTQGRFDTVEY 87  
 Db 14 MTVKVGINGFGRIGRLAFRRRIQDVEGIEVVAINDLTDAKMLAHLLKYDTTQGRFNGTVEV 73  
 QY 88 KEGGFEVNGNFIKVSAERDPENIDWATDGVEIVLEALEGTVKDFGGFDVNGKEIKVSAE 147  
 ; US-09-107-532A-4769



```

QY 445 KI 446
Db 356 NL 357

RESULT 15
US-09-634-238-233
; Sequence 233, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James S.
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross S.
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000-1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
; US-09-634-238-233

```

Query Match 40.3%; Score 918.5; DB 4; Length 340;

Best Local Similarity 46.0%; Pred. No. 2.3e-78; Mismatches 88; Indels 93; Gaps 7;

Matches 1996; Conservative 49; APPLICANT: Glenn, Matthew; APPLICANT: Lubbers, Mark W.; APPLICANT: Dekker, James S.; APPLICANT: Christensson, Anna C.; APPLICANT: Holland, Ross S.; APPLICANT: O'Toole, Paul W.; APPLICANT: Reid, Julian R.; APPLICANT: Coolbear, Timothy; TITLE OF INVENTION: Polynucleotides, materials incorporating them and methods for using them.; FILE REFERENCE: 11000-1043U1; CURRENT APPLICATION NUMBER: US/09/634,238; CURRENT FILING DATE: 2000-08-08; NUMBER OF SEQ ID NOS: 422; SOFTWARE: FastSEQ for Windows Version 4.0; SEQ ID NO 233; LENGTH: 340; TYPE: PRT; ORGANISM: Lactobacillus rhamnosus; US-09-634-238-233

QY 28 MVVKVGINGFGRIGRLAFRRIQNV----EGVEVTRINDLDPNMALAHLLKYDTTQGRFDG 83
Db 1 MTVKIGINGFGRIGRLAFRRIYEIGAKSNDIQVVAINDLTSPTMALAHLLKYDSTHGTFPG 60

QY 84 TVEVKEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGGTVEKDGGFDVNGKFIK 143
Db 61 EVSATDNGIVVDKEYRVYAEPAQONIPW-----VKN----- 92

QY 144 VSAEKDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAIDWATDGVE 203
Db 93 ----- 93

QY 204 IVLEATSSFAKKEAAEKHLHZANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGAS 263
Db 97 YVLECTGPFYTSAAKSQAHDLA-GAKRVLISAPAGK-IKTIIVNVNDDTLNADDKIVSAGS 154

QY 264 CTTNCLAPMAKALHDAGLQKGLMTTHAYTGQMDLDPHRCGDLRRARAGANIVNS 323
Db 155 CTTNCLAPMAYFLNQEGFIEVGTMETTVHAYTSTQMLLDGPVRGGNLRAARSAAANTIPHS 214

QY 324 TGAAKAIGLVTPELNGKLDGAQRVPPGTVLWVTL-DKNSVDEINAAMK--AASN 380
Db 215 TGAAKAIGLVTPELNGKLGQHQAQRVSVDGSLTEVLSILKTKNVTADQVNEAIKKHTENN 274

QY 381 DSFGYTEDPIVSSDIVGVSYGSLSFDATQTKVMEVDGSQLVKVVSWYDNEMSYTAQLVRTL 440
Db 275 PSFGWNDEIVSSDVIGRTYGSIFDPTQTEVTTAGDYQLVKTVAWYDNEYGFTCQMIRTL 334

QY 441 EYFAKI 446
Db 335 LKFATL 340

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:08:58 ; Search time 131 Seconds  
 (without alignments)  
 1.096.710 Million cell updates/sec

Title: US-10-650-369-22  
 Perfect score: 2278  
 Sequence: 1 MKKITGILLLAVILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications AA:  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:  
 4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:  
 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:  
 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:  
 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:  
 RESULT 1  
 US-09-878-766A-22  
 ; Sequence 22, Application US/09878766A  
 ; Patent No. US20020044928A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Potter, Andrew A.  
 ; APPLICANT: Perez-Casal, Jose  
 ; APPLICANT: Fontaine, Michael  
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
 ; AGAINST STREPTOCOCCUS INFECTION  
 ; FILE REFERENCE: 9000-0057  
 ; CURRENT APPLICATION NUMBER: US/09/878,766A  
 ; CURRENT FILING DATE: 2001-09-10  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 22  
 ; LENGTH: 448  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple  
 ; epitope fusion protein

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	9 US-09-878-766A-22	Sequence 22, Appl
2	2278	100.0	448	12 US-10-650-369-22	Sequence 22, Appl
3	1656.5	72.7	336	9 US-09-878-766A-12	Sequence 12, Appl
4	1656.5	72.7	336	10 US-09-878-781-4	Sequence 4, Appl
5	1656.5	72.7	336	12 US-10-650-369-12	Sequence 12, Appl
6	1656.5	72.7	336	14 US-10-134-297-4	Sequence 4, Appl
7	1655.5	72.7	336	12 US-10-282-122A-74379	Sequence 74379, A
8	1650	72.4	335	10 US-09-878-781-14	Sequence 14, Appl
9	1562.5	68.6	336	10 US-09-878-781-16	Sequence 16, Appl
10	1559.5	68.5	336	9 US-09-878-766A-16	Sequence 16, Appl
11	1559.5	68.5	336	10 US-09-878-781-8	Sequence 8, Appl
12	1559.5	68.5	336	12 US-10-650-369-16	Sequence 16, Appl
13	1559.5	68.5	336	14 US-10-134-297-8	Sequence 8, Appl
14	1557.5	68.4	336	9 US-09-878-766A-14	Sequence 14, Appl
15	1557.5	68.4	336	10 US-09-878-781-6	Sequence 6, Appl

Sequence 14, Appl  
 Sequence 6, Appl  
 Sequence 13169, A  
 Sequence 74186, A  
 Sequence 20, Appl  
 Sequence 12, Appl  
 Sequence 20, Appl  
 Sequence 12, Appl  
 Sequence 13593, A  
 Sequence 18, Appl  
 Sequence 10, Appl  
 Sequence 18, Appl  
 Sequence 10, Appl  
 Sequence 72254, A  
 Sequence 54, Appl  
 Sequence 18524, A  
 Sequence 10847, A  
 Sequence 57294, A  
 Sequence 18342, A  
 Sequence 18342, A  
 Sequence 53266, A  
 Sequence 60492, A  
 Sequence 51519, A  
 Sequence 71734, A  
 Sequence 65661, A  
 Sequence 65001, A  
 Sequence 70890, A  
 Sequence 5618, AP  
 Sequence 12540, A  
 Sequence 43997, A  
 Sequence 56828, A

ALIGMENTS

US-09-878-766A-22  
 ; Sequence 22, Application US/09878766A  
 ; Patent No. US20020044928A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Potter, Andrew A.  
 ; APPLICANT: Perez-Casal, Jose  
 ; APPLICANT: Fontaine, Michael  
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN  
 ; AGAINST STREPTOCOCCUS INFECTION  
 ; FILE REFERENCE: 9000-0057  
 ; CURRENT APPLICATION NUMBER: US/09/878,766A  
 ; CURRENT FILING DATE: 2001-09-10  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 22  
 ; LENGTH: 448  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple  
 ; epitope fusion protein

US-09-878-766A-22



; Sequence 4, Application US/09878781  
 ; Publication No. US20030082781A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bolton, Alexandra J.  
 ; APPLICANT: Perez-Casal, Jose  
 ; APPLICANT: Fontaine, Michael  
 ; APPLICANT: Potter, Andrew A.  
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST  
 ; STREPTOCOCCUS INFECTION  
 ; FILE REFERENCE: 9000-0055  
 ; CURRENT FILING DATE: 2002-09-10  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 4  
 ; LENGTH: 336  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus dysgalactiae  
 ; US-09-878-781-4

Query Match 72.7%; Score 1656.5; DB 12; Length 336;  
 Best Local Similarity 79.6%; Pred. No. 3.4e-142;  
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 28 M<sup>Y</sup>V<sup>K</sup>G<sup>I</sup>N<sup>G</sup>F<sup>R</sup>I<sup>G</sup>R<sup>L</sup>A<sup>F</sup>R<sup>R</sup>I<sup>Q</sup>N<sup>V</sup>E<sup>G</sup>V<sup>E</sup>T<sup>R</sup>I<sup>N</sup>D<sup>L</sup>T<sup>D</sup>P<sup>N</sup>M<sup>L</sup>A<sup>H</sup>L<sup>K</sup>Y<sup>D</sup>T<sup>T</sup>Q<sup>G</sup>R<sup>F</sup>D<sup>G</sup>T<sup>V</sup>E<sup>V</sup> 87  
 Db 1 M<sup>Y</sup>V<sup>K</sup>G<sup>I</sup>N<sup>G</sup>F<sup>R</sup>I<sup>G</sup>R<sup>L</sup>A<sup>F</sup>R<sup>R</sup>I<sup>Q</sup>N<sup>V</sup>E<sup>G</sup>V<sup>E</sup>T<sup>R</sup>I<sup>N</sup>D<sup>L</sup>T<sup>D</sup>P<sup>N</sup>M<sup>L</sup>A<sup>H</sup>L<sup>K</sup>Y<sup>D</sup>T<sup>T</sup>Q<sup>G</sup>R<sup>F</sup>D<sup>G</sup>T<sup>V</sup>E<sup>V</sup> 60

Qy 88 K<sup>E</sup>GG<sup>F</sup>E<sup>V</sup>N<sup>G</sup>N<sup>F</sup>I<sup>K</sup>Y<sup>S</sup>A<sup>E</sup>R<sup>D</sup>P<sup>E</sup>N<sup>I</sup>D<sup>W</sup>A<sup>T</sup>D<sup>G</sup>V<sup>E</sup>I<sup>V</sup>L<sup>E</sup>A<sup>L</sup>E<sup>G</sup>T<sup>V</sup>E<sup>V</sup>K<sup>D</sup>G<sup>G</sup>F<sup>D</sup>V<sup>N</sup>G<sup>K</sup>F<sup>I</sup>K<sup>V</sup>S<sup>A</sup>E 147  
 Db 61 K<sup>E</sup>GG<sup>F</sup>E<sup>V</sup>N<sup>G</sup>N<sup>F</sup>I<sup>K</sup>Y<sup>S</sup>A<sup>E</sup>R<sup>D</sup>P<sup>E</sup>-----81

Qy 148 K<sup>D</sup>P<sup>E</sup>Q<sup>I</sup>D<sup>W</sup>A<sup>T</sup>D<sup>G</sup>V<sup>E</sup>I<sup>V</sup>L<sup>E</sup>I<sup>D</sup>G<sup>T</sup>V<sup>E</sup>V<sup>K</sup>E<sup>G</sup>G<sup>F</sup>E<sup>V</sup>N<sup>G</sup>Q<sup>F</sup>V<sup>K</sup>V<sup>S</sup>A<sup>E</sup>R<sup>P</sup>A<sup>N</sup>I<sup>D</sup>W<sup>A</sup>T<sup>D</sup>G<sup>V</sup>E<sup>I</sup>V<sup>L</sup>E 207  
 Db 82 -----N<sup>I</sup>D<sup>W</sup>A<sup>T</sup>D<sup>G</sup>V<sup>E</sup>I<sup>V</sup>L<sup>E</sup>-----95

Qy 208 A<sup>T</sup>S<sup>F</sup>F<sup>A</sup>K<sup>E</sup>A<sup>A</sup>E<sup>K</sup>H<sup>L</sup>H<sup>A</sup>N<sup>G</sup>A<sup>K</sup>K<sup>V</sup>V<sup>I</sup>T<sup>A</sup>P<sup>G</sup>G<sup>N</sup>D<sup>V</sup>K<sup>T</sup>V<sup>V</sup>F<sup>N</sup>T<sup>N</sup>H<sup>D</sup>I<sup>L</sup>D<sup>G</sup>T<sup>E</sup>V<sup>I</sup>S<sup>G</sup>A<sup>S</sup>C<sup>T</sup>T<sup>N</sup> 267  
 Db 96 A<sup>T</sup>G<sup>F</sup>F<sup>A</sup>K<sup>E</sup>A<sup>A</sup>E<sup>K</sup>H<sup>L</sup>H<sup>A</sup>N<sup>G</sup>A<sup>K</sup>K<sup>V</sup>V<sup>I</sup>T<sup>A</sup>P<sup>G</sup>G<sup>N</sup>D<sup>V</sup>K<sup>T</sup>V<sup>V</sup>F<sup>N</sup>T<sup>N</sup>H<sup>D</sup>I<sup>L</sup>D<sup>G</sup>T<sup>E</sup>V<sup>I</sup>S<sup>G</sup>A<sup>S</sup>C<sup>T</sup>T<sup>N</sup> 155

Qy 268 C<sup>L</sup>A<sup>P</sup>M<sup>A</sup>K<sup>A</sup>L<sup>H</sup>D<sup>A</sup>F<sup>G</sup>I<sup>Q</sup>K<sup>G</sup>L<sup>M</sup>T<sup>T</sup>I<sup>H</sup>A<sup>Y</sup>T<sup>G</sup>D<sup>Q</sup>M<sup>I</sup>L<sup>D</sup>G<sup>P</sup>H<sup>R</sup>G<sup>G</sup>D<sup>L</sup>R<sup>R</sup>A<sup>R</sup>G<sup>A</sup>A<sup>N</sup>I<sup>V</sup>P<sup>N</sup>S<sup>T</sup>G<sup>A</sup>A 327  
 Db 156 C<sup>L</sup>A<sup>P</sup>M<sup>A</sup>K<sup>A</sup>L<sup>H</sup>D<sup>A</sup>F<sup>G</sup>I<sup>Q</sup>K<sup>G</sup>L<sup>M</sup>T<sup>T</sup>I<sup>H</sup>A<sup>Y</sup>T<sup>G</sup>D<sup>Q</sup>M<sup>I</sup>L<sup>D</sup>G<sup>P</sup>H<sup>R</sup>G<sup>G</sup>D<sup>L</sup>R<sup>R</sup>A<sup>R</sup>G<sup>A</sup>A<sup>N</sup>I<sup>V</sup>P<sup>N</sup>S<sup>T</sup>G<sup>A</sup>A 215

Qy 328 K<sup>A</sup>I<sup>G</sup>L<sup>V</sup>I<sup>P</sup>E<sup>L</sup>N<sup>G</sup>K<sup>L</sup>D<sup>G</sup>A<sup>A</sup>Q<sup>R</sup>V<sup>P</sup>P<sup>V</sup>T<sup>G</sup>S<sup>V</sup>T<sup>E</sup>L<sup>V</sup>V<sup>T</sup>L<sup>D</sup>K<sup>N</sup>V<sup>S</sup>V<sup>D</sup>E<sup>I</sup>N<sup>A</sup>A<sup>M</sup>K<sup>A</sup>A<sup>S</sup>N<sup>D</sup>S<sup>F</sup>G<sup>Y</sup>T<sup>E</sup> 387  
 Db 216 K<sup>A</sup>I<sup>G</sup>L<sup>V</sup>I<sup>P</sup>E<sup>L</sup>N<sup>G</sup>K<sup>L</sup>D<sup>G</sup>A<sup>A</sup>Q<sup>R</sup>V<sup>P</sup>P<sup>V</sup>T<sup>G</sup>S<sup>V</sup>T<sup>E</sup>L<sup>V</sup>V<sup>T</sup>L<sup>D</sup>K<sup>N</sup>V<sup>S</sup>V<sup>D</sup>E<sup>I</sup>N<sup>A</sup>A<sup>M</sup>K<sup>A</sup>A<sup>S</sup>N<sup>D</sup>S<sup>F</sup>G<sup>Y</sup>T<sup>E</sup> 275

Qy 148 K<sup>D</sup>P<sup>E</sup>Q<sup>I</sup>D<sup>W</sup>A<sup>T</sup>D<sup>G</sup>V<sup>E</sup>I<sup>V</sup>L<sup>E</sup>I<sup>D</sup>G<sup>T</sup>V<sup>E</sup>V<sup>K</sup>E<sup>G</sup>G<sup>F</sup>E<sup>V</sup>N<sup>G</sup>Q<sup>F</sup>V<sup>K</sup>V<sup>S</sup>A<sup>E</sup>R<sup>P</sup>A<sup>N</sup>I<sup>D</sup>W<sup>A</sup>T<sup>D</sup>G<sup>V</sup>E<sup>I</sup>V<sup>L</sup>E 207  
 Db 96 A<sup>T</sup>G<sup>F</sup>F<sup>A</sup>K<sup>E</sup>A<sup>A</sup>E<sup>K</sup>H<sup>L</sup>H<sup>A</sup>N<sup>G</sup>A<sup>K</sup>K<sup>V</sup>V<sup>I</sup>T<sup>A</sup>P<sup>G</sup>G<sup>N</sup>D<sup>V</sup>K<sup>T</sup>V<sup>V</sup>F<sup>N</sup>T<sup>N</sup>H<sup>D</sup>I<sup>L</sup>D<sup>G</sup>T<sup>E</sup>V<sup>I</sup>S<sup>G</sup>A<sup>S</sup>C<sup>T</sup>T<sup>N</sup> 155

Db 82 -----N<sup>I</sup>D<sup>W</sup>A<sup>T</sup>D<sup>G</sup>V<sup>E</sup>I<sup>V</sup>L<sup>E</sup>-----95

Qy 268 C<sup>L</sup>A<sup>P</sup>M<sup>A</sup>K<sup>A</sup>L<sup>H</sup>D<sup>A</sup>F<sup>G</sup>I<sup>Q</sup>K<sup>G</sup>L<sup>M</sup>T<sup>T</sup>I<sup>H</sup>A<sup>Y</sup>T<sup>G</sup>D<sup>Q</sup>M<sup>I</sup>L<sup>D</sup>G<sup>P</sup>H<sup>R</sup>G<sup>G</sup>D<sup>L</sup>R<sup>R</sup>A<sup>R</sup>G<sup>A</sup>A<sup>N</sup>I<sup>V</sup>P<sup>N</sup>S<sup>T</sup>G<sup>A</sup>A 327  
 Db 156 C<sup>L</sup>A<sup>P</sup>M<sup>A</sup>K<sup>A</sup>L<sup>H</sup>D<sup>A</sup>F<sup>G</sup>I<sup>Q</sup>K<sup>G</sup>L<sup>M</sup>T<sup>T</sup>I<sup>H</sup>A<sup>Y</sup>T<sup>G</sup>D<sup>Q</sup>M<sup>I</sup>L<sup>D</sup>G<sup>P</sup>H<sup>R</sup>G<sup>G</sup>D<sup>L</sup>R<sup>R</sup>A<sup>R</sup>G<sup>A</sup>A<sup>N</sup>I<sup>V</sup>P<sup>N</sup>S<sup>T</sup>G<sup>A</sup>A 215

Qy 328 K<sup>A</sup>I<sup>G</sup>L<sup>V</sup>I<sup>P</sup>E<sup>L</sup>N<sup>G</sup>K<sup>L</sup>D<sup>G</sup>A<sup>A</sup>Q<sup>R</sup>V<sup>P</sup>P<sup>V</sup>T<sup>G</sup>S<sup>V</sup>T<sup>E</sup>L<sup>V</sup>V<sup>T</sup>L<sup>D</sup>K<sup>N</sup>V<sup>S</sup>V<sup>D</sup>E<sup>I</sup>N<sup>A</sup>A<sup>M</sup>K<sup>A</sup>A<sup>S</sup>N<sup>D</sup>S<sup>F</sup>G<sup>Y</sup>T<sup>E</sup> 387  
 Db 216 K<sup>A</sup>I<sup>G</sup>L<sup>V</sup>I<sup>P</sup>E<sup>L</sup>N<sup>G</sup>K<sup>L</sup>D<sup>G</sup>A<sup>A</sup>Q<sup>R</sup>V<sup>P</sup>P<sup>V</sup>T<sup>G</sup>S<sup>V</sup>T<sup>E</sup>L<sup>V</sup>V<sup>T</sup>L<sup>D</sup>K<sup>N</sup>V<sup>S</sup>V<sup>D</sup>E<sup>I</sup>N<sup>A</sup>A<sup>M</sup>K<sup>A</sup>A<sup>S</sup>N<sup>D</sup>S<sup>F</sup>G<sup>Y</sup>T<sup>E</sup> 275

Qy 388 D<sup>P</sup>I<sup>V</sup>S<sup>S</sup>D<sup>I</sup>V<sup>G</sup>V<sup>S</sup>G<sup>I</sup>F<sup>D</sup>A<sup>T</sup>Q<sup>T</sup>K<sup>V</sup>M<sup>E</sup>V<sup>D</sup>G<sup>S</sup>Q<sup>L</sup>V<sup>K</sup>V<sup>S</sup>W<sup>Y</sup>D<sup>N</sup>E<sup>M</sup>S<sup>T</sup>A<sup>Q</sup>L<sup>V</sup>R<sup>T</sup>L<sup>E</sup>Y<sup>F</sup>A<sup>K</sup>I<sup>A</sup> 447  
 Db 276 D<sup>P</sup>I<sup>V</sup>S<sup>S</sup>D<sup>I</sup>V<sup>G</sup>V<sup>S</sup>G<sup>I</sup>F<sup>D</sup>A<sup>T</sup>Q<sup>T</sup>K<sup>V</sup>M<sup>E</sup>V<sup>D</sup>G<sup>S</sup>Q<sup>L</sup>V<sup>K</sup>V<sup>S</sup>W<sup>Y</sup>D<sup>N</sup>E<sup>M</sup>S<sup>T</sup>A<sup>Q</sup>L<sup>V</sup>R<sup>T</sup>L<sup>E</sup>Y<sup>F</sup>A<sup>K</sup>I<sup>A</sup> 335

Qy 448 K 448  
 Db 336 K 336

RESULT 6  
 US-10-134-297-4  
 ; Sequence 4, Application US/10134297  
 ; Publication No. US20030165524A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bolton, Alexandra J.  
 ; APPLICANT: Perez-Casal, Jose  
 ; APPLICANT: Fontaine, Michael  
 ; APPLICANT: Potter, Andrew A.  
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST  
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN  
 ; FILE REFERENCE: 9000-0055-20  
 ; CURRENT APPLICATION NUMBER: US/10/134-297  
 ; CURRENT FILING DATE: 2002-04-26  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SEQ ID NO 4  
 ; LENGTH: 336  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus dysgalactiae  
 US-10-134-297-4

Query Match 72.7%; Score 1656.5; DB 14; Length 336;  
 Best Local Similarity 79.6%; Pred. No. 3.4e-142;  
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 28 M<sup>Y</sup>V<sup>K</sup>G<sup>I</sup>N<sup>G</sup>F<sup>R</sup>I<sup>G</sup>R<sup>L</sup>A<sup>F</sup>R<sup>R</sup>I<sup>Q</sup>N<sup>V</sup>E<sup>G</sup>V<sup>E</sup>T<sup>R</sup>I<sup>N</sup>D<sup>L</sup>T<sup>D</sup>P<sup>N</sup>M<sup>L</sup>A<sup>H</sup>L<sup>K</sup>Y<sup>D</sup>T<sup>T</sup>Q<sup>G</sup>R<sup>F</sup>D<sup>G</sup>T<sup>V</sup>E<sup>V</sup> 87  
 Db 1 M<sup>Y</sup>V<sup>K</sup>G<sup>I</sup>N<sup>G</sup>F<sup>R</sup>I<sup>G</sup>R<sup>L</sup>A<sup>F</sup>R<sup>R</sup>I<sup>Q</sup>N<sup>V</sup>E<sup>G</sup>V<sup>E</sup>T<sup>R</sup>I<sup>N</sup>D<sup>L</sup>T<sup>D</sup>P<sup>N</sup>M<sup>L</sup>A<sup>H</sup>L<sup>K</sup>Y<sup>D</sup>T<sup>T</sup>Q<sup>G</sup>R<sup>F</sup>D<sup>G</sup>T<sup>V</sup>E<sup>V</sup> 60

1 SOFTWARE: PatentIn version 3.1  
 1 ; SEQ ID NO 74379  
 1 ; LENGTH: 336  
 1 ; TYPE: PRT  
 1 ; ORGANISM: *Streptococcus pyogenes*  
 1 ; US-10-282-122A-74379

2 Query Match  
 2 Best Local Similarity 79.3%; Pred. No. 4.e-142;  
 2 Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

3 QY 28 M<sub>1</sub>V<sub>2</sub>K<sub>3</sub>V<sub>4</sub>G<sub>5</sub>I<sub>6</sub>N<sub>7</sub>G<sub>8</sub>F<sub>9</sub>R<sub>10</sub>R<sub>11</sub>I<sub>12</sub>R<sub>13</sub>I<sub>14</sub>R<sub>15</sub>I<sub>16</sub>K<sub>17</sub>D<sub>18</sub>T<sub>19</sub>G<sub>20</sub>T<sub>21</sub>E<sub>22</sub>V<sub>23</sub>G<sub>24</sub>V<sub>25</sub>E<sub>26</sub>I<sub>27</sub>V<sub>28</sub>E<sub>29</sub>A<sub>30</sub>E<sub>31</sub>R<sub>32</sub>P<sub>33</sub>E<sub>34</sub>N<sub>35</sub>I<sub>36</sub>R<sub>37</sub>A<sub>38</sub>G<sub>39</sub>A<sub>40</sub>A<sub>41</sub>N<sub>42</sub>I<sub>43</sub>A<sub>44</sub>N<sub>45</sub>I<sub>46</sub>V<sub>47</sub>E<sub>48</sub>I<sub>49</sub>V<sub>50</sub>E<sub>51</sub>A<sub>52</sub>E<sub>53</sub>R<sub>54</sub>I<sub>55</sub>V<sub>56</sub>A<sub>57</sub>E<sub>58</sub>I<sub>59</sub>R<sub>60</sub>A<sub>61</sub>G<sub>62</sub>F<sub>63</sub>R<sub>64</sub>R<sub>65</sub>I<sub>66</sub>R<sub>67</sub>A<sub>68</sub>G<sub>69</sub>A<sub>70</sub>A<sub>71</sub>N<sub>72</sub>I<sub>73</sub>A<sub>74</sub>N<sub>75</sub>I<sub>76</sub>R<sub>77</sub>A<sub>78</sub>G<sub>79</sub>A<sub>80</sub>A<sub>81</sub>N<sub>82</sub>I<sub>83</sub>R<sub>84</sub>A<sub>85</sub>G<sub>86</sub>F<sub>87</sub>R<sub>88</sub>R<sub>89</sub>I<sub>90</sub>R<sub>91</sub>A<sub>92</sub>G<sub>93</sub>A<sub>94</sub>A<sub>95</sub>I<sub>96</sub>R<sub>97</sub>A<sub>98</sub>G<sub>99</sub>A<sub>100</sub>A<sub>101</sub>N<sub>102</sub>I<sub>103</sub>R<sub>104</sub>A<sub>105</sub>G<sub>106</sub>F<sub>107</sub>R<sub>108</sub>R<sub>109</sub>I<sub>110</sub>R<sub>111</sub>A<sub>112</sub>G<sub>113</sub>A<sub>114</sub>A<sub>115</sub>N<sub>116</sub>I<sub>117</sub>R<sub>118</sub>A<sub>119</sub>G<sub>120</sub>F<sub>121</sub>R<sub>122</sub>R<sub>123</sub>I<sub>124</sub>R<sub>125</sub>A<sub>126</sub>G<sub>127</sub>A<sub>128</sub>A<sub>129</sub>N<sub>130</sub>I<sub>131</sub>R<sub>132</sub>A<sub>133</sub>G<sub>134</sub>F<sub>135</sub>R<sub>136</sub>R<sub>137</sub>I<sub>138</sub>R<sub>139</sub>A<sub>140</sub>G<sub>141</sub>A<sub>142</sub>A<sub>143</sub>N<sub>144</sub>I<sub>145</sub>R<sub>146</sub>A<sub>147</sub>G<sub>148</sub>F<sub>149</sub>R<sub>150</sub>R<sub>151</sub>I<sub>152</sub>R<sub>153</sub>A<sub>154</sub>G<sub>155</sub>A<sub>156</sub>A<sub>157</sub>N<sub>158</sub>I<sub>159</sub>R<sub>160</sub>A<sub>161</sub>G<sub>162</sub>F<sub>163</sub>R<sub>164</sub>R<sub>165</sub>I<sub>166</sub>R<sub>167</sub>A<sub>168</sub>G<sub>169</sub>A<sub>170</sub>A<sub>171</sub>N<sub>172</sub>I<sub>173</sub>R<sub>174</sub>A<sub>175</sub>G<sub>176</sub>F<sub>177</sub>R<sub>178</sub>R<sub>179</sub>I<sub>180</sub>R<sub>181</sub>A<sub>182</sub>G<sub>183</sub>A<sub>184</sub>A<sub>185</sub>N<sub>186</sub>I<sub>187</sub>R<sub>188</sub>A<sub>189</sub>G<sub>190</sub>F<sub>191</sub>R<sub>192</sub>R<sub>193</sub>I<sub>194</sub>R<sub>195</sub>A<sub>196</sub>G<sub>197</sub>A<sub>198</sub>A<sub>199</sub>N<sub>200</sub>I<sub>201</sub>R<sub>202</sub>A<sub>203</sub>G<sub>204</sub>F<sub>205</sub>R<sub>206</sub>R<sub>207</sub>I<sub>208</sub>R<sub>209</sub>A<sub>210</sub>G<sub>211</sub>A<sub>212</sub>A<sub>213</sub>N<sub>214</sub>I<sub>215</sub>R<sub>216</sub>A<sub>217</sub>G<sub>218</sub>F<sub>219</sub>R<sub>220</sub>R<sub>221</sub>I<sub>222</sub>R<sub>223</sub>A<sub>224</sub>G<sub>225</sub>A<sub>226</sub>A<sub>227</sub>N<sub>228</sub>I<sub>229</sub>R<sub>230</sub>A<sub>231</sub>G<sub>232</sub>F<sub>233</sub>R<sub>234</sub>R<sub>235</sub>I<sub>236</sub>R<sub>237</sub>A<sub>238</sub>G<sub>239</sub>A<sub>240</sub>A<sub>241</sub>N<sub>242</sub>I<sub>243</sub>R<sub>244</sub>A<sub>245</sub>G<sub>246</sub>F<sub>247</sub>R<sub>248</sub>R<sub>249</sub>I<sub>250</sub>R<sub>251</sub>A<sub>252</sub>G<sub>253</sub>A<sub>254</sub>A<sub>255</sub>N<sub>256</sub>I<sub>257</sub>R<sub>258</sub>A<sub>259</sub>G<sub>260</sub>F<sub>261</sub>R<sub>262</sub>R<sub>263</sub>I<sub>264</sub>R<sub>265</sub>A<sub>266</sub>G<sub>267</sub>A<sub>268</sub>A<sub>269</sub>N<sub>270</sub>I<sub>271</sub>R<sub>272</sub>A<sub>273</sub>G<sub>274</sub>F<sub>275</sub>R<sub>276</sub>R<sub>277</sub>I<sub>278</sub>R<sub>279</sub>A<sub>280</sub>G<sub>281</sub>A<sub>282</sub>A<sub>283</sub>N<sub>284</sub>I<sub>285</sub>R<sub>286</sub>A<sub>287</sub>G<sub>288</sub>F<sub>289</sub>R<sub>290</sub>R<sub>291</sub>I<sub>292</sub>R<sub>293</sub>A<sub>294</sub>G<sub>295</sub>A<sub>296</sub>A<sub>297</sub>N<sub>298</sub>I<sub>299</sub>R<sub>300</sub>A<sub>301</sub>G<sub>302</sub>F<sub>303</sub>R<sub>304</sub>R<sub>305</sub>I<sub>306</sub>R<sub>307</sub>A<sub>308</sub>G<sub>309</sub>A<sub>310</sub>A<sub>311</sub>N<sub>312</sub>I<sub>313</sub>R<sub>314</sub>A<sub>315</sub>G<sub>316</sub>F<sub>317</sub>R<sub>318</sub>R<sub>319</sub>I<sub>320</sub>R<sub>321</sub>A<sub>322</sub>G<sub>323</sub>A<sub>324</sub>A<sub>325</sub>N<sub>326</sub>I<sub>327</sub>R<sub>328</sub>A<sub>329</sub>G<sub>330</sub>F<sub>331</sub>R<sub>332</sub>R<sub>333</sub>I<sub>334</sub>R<sub>335</sub>A<sub>336</sub>G<sub>337</sub>A<sub>338</sub>A<sub>339</sub>N<sub>340</sub>I<sub>341</sub>R<sub>342</sub>A<sub>343</sub>G<sub>344</sub>F<sub>345</sub>R<sub>346</sub>R<sub>347</sub>I<sub>348</sub>R<sub>349</sub>A<sub>350</sub>G<sub>351</sub>A<sub>352</sub>A<sub>353</sub>N<sub>354</sub>I<sub>355</sub>R<sub>356</sub>A<sub>357</sub>G<sub>358</sub>F<sub>359</sub>R<sub>360</sub>R<sub>361</sub>I<sub>362</sub>R<sub>363</sub>A<sub>364</sub>G<sub>365</sub>A<sub>366</sub>A<sub>367</sub>N<sub>368</sub>I<sub>369</sub>R<sub>370</sub>A<sub>371</sub>G<sub>372</sub>F<sub>373</sub>R<sub>374</sub>R<sub>375</sub>I<sub>376</sub>R<sub>377</sub>A<sub>378</sub>G<sub>379</sub>A<sub>380</sub>A<sub>381</sub>N<sub>382</sub>I<sub>383</sub>R<sub>384</sub>A<sub>385</sub>G<sub>386</sub>F<sub>387</sub>R<sub>388</sub>R<sub>389</sub>I<sub>390</sub>R<sub>391</sub>A<sub>392</sub>G<sub>393</sub>A<sub>394</sub>A<sub>395</sub>N<sub>396</sub>I<sub>397</sub>R<sub>398</sub>A<sub>399</sub>G<sub>400</sub>F<sub>401</sub>R<sub>402</sub>R<sub>403</sub>I<sub>404</sub>R<sub>405</sub>A<sub>406</sub>G<sub>407</sub>A<sub>408</sub>A<sub>409</sub>N<sub>410</sub>I<sub>411</sub>R<sub>412</sub>A<sub>413</sub>G<sub>414</sub>F<sub>415</sub>R<sub>416</sub>R<sub>417</sub>I<sub>418</sub>R<sub>419</sub>A<sub>420</sub>G<sub>421</sub>A<sub>422</sub>A<sub>423</sub>N<sub>424</sub>I<sub>425</sub>R<sub>426</sub>A<sub>427</sub>G<sub>428</sub>F<sub>429</sub>R<sub>430</sub>R<sub>431</sub>I<sub>432</sub>R<sub>433</sub>A<sub>434</sub>G<sub>435</sub>A<sub>436</sub>A<sub>437</sub>N<sub>438</sub>I<sub>439</sub>R<sub>440</sub>A<sub>441</sub>G<sub>442</sub>F<sub>443</sub>R<sub>444</sub>R<sub>445</sub>I<sub>446</sub>R<sub>447</sub>A<sub>448</sub>G<sub>449</sub>A<sub>450</sub>A<sub>451</sub>N<sub>452</sub>I<sub>453</sub>R<sub>454</sub>A<sub>455</sub>G<sub>456</sub>F<sub>457</sub>R<sub>458</sub>R<sub>459</sub>I<sub>460</sub>R<sub>461</sub>A<sub>462</sub>G<sub>463</sub>A<sub>464</sub>A<sub>465</sub>N<sub>466</sub>I<sub>467</sub>R<sub>468</sub>A<sub>469</sub>G<sub>470</sub>F<sub>471</sub>R<sub>472</sub>R<sub>473</sub>I<sub>474</sub>R<sub>475</sub>A<sub>476</sub>G<sub>477</sub>A<sub>478</sub>A<sub>479</sub>N<sub>480</sub>I<sub>481</sub>R<sub>482</sub>A<sub>483</sub>G<sub>484</sub>F<sub>485</sub>R<sub>486</sub>R<sub>487</sub>I<sub>488</sub>R<sub>489</sub>A<sub>490</sub>G<sub>491</sub>A<sub>492</sub>A<sub>493</sub>N<sub>494</sub>I<sub>495</sub>R<sub>496</sub>A<sub>497</sub>G<sub>498</sub>F<sub>499</sub>R<sub>500</sub>R<sub>501</sub>I<sub>502</sub>R<sub>503</sub>A<sub>504</sub>G<sub>505</sub>A<sub>506</sub>A<sub>507</sub>N<sub>508</sub>I<sub>509</sub>R<sub>510</sub>A<sub>511</sub>G<sub>512</sub>F<sub>513</sub>R<sub>514</sub>R<sub>515</sub>I<sub>516</sub>R<sub>517</sub>A<sub>518</sub>G<sub>519</sub>A<sub>520</sub>A<sub>521</sub>N<sub>522</sub>I<sub>523</sub>R<sub>524</sub>A<sub>525</sub>G<sub>526</sub>F<sub>527</sub>R<sub>528</sub>R<sub>529</sub>I<sub>530</sub>R<sub>531</sub>A<sub>532</sub>G<sub>533</sub>A<sub>534</sub>A<sub>535</sub>N<sub>536</sub>I<sub>537</sub>R<sub>538</sub>A<sub>539</sub>G<sub>540</sub>F<sub>541</sub>R<sub>542</sub>R<sub>543</sub>I<sub>544</sub>R<sub>545</sub>A<sub>546</sub>G<sub>547</sub>A<sub>548</sub>A<sub>549</sub>N<sub>550</sub>I<sub>551</sub>R<sub>552</sub>A<sub>553</sub>G<sub>554</sub>F<sub>555</sub>R<sub>556</sub>R<sub>557</sub>I<sub>558</sub>R<sub>559</sub>A<sub>560</sub>G<sub>561</sub>A<sub>562</sub>A<sub>563</sub>N<sub>564</sub>I<sub>565</sub>R<sub>566</sub>A<sub>567</sub>G<sub>568</sub>F<sub>569</sub>R<sub>570</sub>R<sub>571</sub>I<sub>572</sub>R<sub>573</sub>A<sub>574</sub>G<sub>575</sub>A<sub>576</sub>A<sub>577</sub>N<sub>578</sub>I<sub>579</sub>R<sub>580</sub>A<sub>581</sub>G<sub>582</sub>F<sub>583</sub>R<sub>584</sub>R<sub>585</sub>I<sub>586</sub>R<sub>587</sub>A<sub>588</sub>G<sub>589</sub>A<sub>590</sub>A<sub>591</sub>N<sub>592</sub>I<sub>593</sub>R<sub>594</sub>A<sub>595</sub>G<sub>596</sub>F<sub>597</sub>R<sub>598</sub>R<sub>599</sub>I<sub>600</sub>R<sub>601</sub>A<sub>602</sub>G<sub>603</sub>A<sub>604</sub>A<sub>605</sub>N<sub>606</sub>I<sub>607</sub>R<sub>608</sub>A<sub>609</sub>G<sub>610</sub>F<sub>611</sub>R<sub>612</sub>R<sub>613</sub>I<sub>614</sub>R<sub>615</sub>A<sub>616</sub>G<sub>617</sub>A<sub>618</sub>A<sub>619</sub>N<sub>620</sub>I<sub>621</sub>R<sub>622</sub>A<sub>623</sub>G<sub>624</sub>F<sub>625</sub>R<sub>626</sub>R<sub>627</sub>I<sub>628</sub>R<sub>629</sub>A<sub>630</sub>G<sub>631</sub>A<sub>632</sub>A<sub>633</sub>N<sub>634</sub>I<sub>635</sub>R<sub>636</sub>A<sub>637</sub>G<sub>638</sub>F<sub>639</sub>R<sub>640</sub>R<sub>641</sub>I<sub>642</sub>R<sub>643</sub>A<sub>644</sub>G<sub>645</sub>A<sub>646</sub>A<sub>647</sub>N<sub>648</sub>I<sub>649</sub>R<sub>650</sub>A<sub>651</sub>G<sub>652</sub>F<sub>653</sub>R<sub>654</sub>R<sub>655</sub>I<sub>656</sub>R<sub>657</sub>A<sub>658</sub>G<sub>659</sub>A<sub>660</sub>A<sub>661</sub>N<sub>662</sub>I<sub>663</sub>R<sub>664</sub>A<sub>665</sub>G<sub>666</sub>F<sub>667</sub>R<sub>668</sub>R<sub>669</sub>I<sub>670</sub>R<sub>671</sub>A<sub>672</sub>G<sub>673</sub>A<sub>674</sub>A<sub>675</sub>N<sub>676</sub>I<sub>677</sub>R<sub>678</sub>A<sub>679</sub>G<sub>680</sub>F<sub>681</sub>R<sub>682</sub>R<sub>683</sub>I<sub>684</sub>R<sub>685</sub>A<sub>686</sub>G<sub>687</sub>A<sub>688</sub>A<sub>689</sub>N<sub>690</sub>I<sub>691</sub>R<sub>692</sub>A<sub>693</sub>G<sub>694</sub>F<sub>695</sub>R<sub>696</sub>R<sub>697</sub>I<sub>698</sub>R<sub>699</sub>A<sub>700</sub>G<sub>701</sub>A<sub>702</sub>A<sub>703</sub>N<sub>704</sub>I<sub>705</sub>R<sub>706</sub>A<sub>707</sub>G<sub>708</sub>F<sub>709</sub>R<sub>710</sub>R<sub>711</sub>I<sub>712</sub>R<sub>713</sub>A<sub>714</sub>G<sub>715</sub>A<sub>716</sub>A<sub>717</sub>N<sub>718</sub>I<sub>719</sub>R<sub>720</sub>A<sub>721</sub>G<sub>722</sub>F<sub>723</sub>R<sub>724</sub>R<sub>725</sub>I<sub>726</sub>R<sub>727</sub>A<sub>728</sub>G<sub>729</sub>A<sub>730</sub>A<sub>731</sub>N<sub>732</sub>I<sub>733</sub>R<sub>734</sub>A<sub>735</sub>G<sub>736</sub>F<sub>737</sub>R<sub>738</sub>R<sub>739</sub>I<sub>740</sub>R<sub>741</sub>A<sub>742</sub>G<sub>743</sub>A<sub>744</sub>A<sub>745</sub>N<sub>746</sub>I<sub>747</sub>R<sub>748</sub>A<sub>749</sub>G<sub>750</sub>F<sub>751</sub>R<sub>752</sub>R<sub>753</sub>I<sub>754</sub>R<sub>755</sub>A<sub>756</sub>G<sub>757</sub>A<sub>758</sub>A<sub>759</sub>N<sub>760</sub>I<sub>761</sub>R<sub>762</sub>A<sub>763</sub>G<sub>764</sub>F<sub>765</sub>R<sub>766</sub>R<sub>767</sub>I<sub>768</sub>R<sub>769</sub>A<sub>770</sub>G<sub>771</sub>A<sub>772</sub>A<sub>773</sub>N<sub>774</sub>I<sub>775</sub>R<sub>776</sub>A<sub>777</sub>G<sub>778</sub>F<sub>779</sub>R<sub>780</sub>R<sub>781</sub>I<sub>782</sub>R<sub>783</sub>A<sub>784</sub>G<sub>785</sub>A<sub>786</sub>A<sub>787</sub>N<sub>788</sub>I<sub>789</sub>R<sub>790</sub>A<sub>791</sub>G<sub>792</sub>F<sub>793</sub>R<sub>794</sub>R<sub>795</sub>I<sub>796</sub>R<sub>797</sub>A<sub>798</sub>G<sub>799</sub>A<sub>800</sub>A<sub>801</sub>N<sub>802</sub>I<sub>803</sub>R<sub>804</sub>A<sub>805</sub>G<sub>806</sub>F<sub>807</sub>R<sub>808</sub>R<sub>809</sub>I<sub>810</sub>R<sub>811</sub>A<sub>812</sub>G<sub>813</sub>A<sub>814</sub>A<sub>815</sub>N<sub>816</sub>I<sub>817</sub>R<sub>818</sub>A<sub>819</sub>G<sub>820</sub>F<sub>821</sub>R<sub>822</sub>R<sub>823</sub>I<sub>824</sub>R<sub>825</sub>A<sub>826</sub>G<sub>827</sub>A<sub>828</sub>A<sub>829</sub>N<sub>830</sub>I<sub>831</sub>R<sub>832</sub>A<sub>833</sub>G<sub>834</sub>F<sub>835</sub>R<sub>836</sub>R<sub>837</sub>I<sub>838</sub>R<sub>839</sub>A<sub>840</sub>G<sub>841</sub>A<sub>842</sub>A<sub>843</sub>A<sub>844</sub>N<sub>845</sub>I<sub>846</sub>R<sub>847</sub>A<sub>848</sub>G<sub>849</sub>F<sub>850</sub>R<sub>851</sub>R<sub>852</sub>I<sub>853</sub>R<sub>854</sub>A<sub>855</sub>G<sub>856</sub>A<sub>857</sub>A<sub>858</sub>N<sub>859</sub>I<sub>860</sub>R<sub>861</sub>A<sub>862</sub>G<sub>863</sub>F<sub>864</sub>R<sub>865</sub>R<sub>866</</sub>



Db 276 DPIVSSDIIGMAYGSLFDATQTKVQTVDDGNQLVKVTVSWYDNEMSYTAQLVRTLEYFAKIA 335

Qy 448 K 448

Db 336 K 336

RESULT 11  
US-09-878-781-8

; Sequence 8, Application US/09878781  
; Publication No. US20030082781A1  
; GENERAL INFORMATION:  
; APPLICANT: Bolton, Michael J.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael A.  
; APPLICANT: Potter, Andrew A.  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GpC PROTEIN AGAINST  
; STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0055  
; CURRENT APPLICATION NUMBER: US/09/878,781  
; CURRENT FILING DATE: 2002-09-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus uberis  
; US-09-878-781-8

Query Match 68.5%; Score 1559.5; DB 10; Length 336;  
Best Local Similarity 73.9%; Pred. No. 2.3e-133;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

Qy 28 MVVKVGINGFGRIGRLAARRIIONVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVV 87  
Db 1 MVVKVGINGFGRIGRLAARRIIONVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVV 60

Qy 88 KEGGFEVNGNFIKVSAAEDPENIDWATDGVETVLEALEGTVEVKDGGFDVNGKPIKVSAE 147  
Db 61 KDGGEVNGNFIKVSAAEKDPE-----  
----- 81

Qy 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREBAPIIDWATDGVETVLE 207  
Db 82 -----  
----- 95

Qy 208 ATSFIAKKEAAEKLHANGAKKVITAPGGNDVKTIVFNTNHDILDGTETVLSGASCTTN 267  
Db 96 ATGFFAKKAAAEEKLHANGAKKVITAPGGDDVKTIVFNTNHDILDGTETVLSGASCTTN 155

Qy 268 CLAPMAKALHDAFGIQKGLMTTIIHAYTGQMDLGPGRGGDLRRARAAGAAANIVPNSTGAA 327  
Db 156 CLAPMAKALQDNFGVKQGLMTTIIHAYTGQMDLGPGRGGDLRRARAAGASNIVPNSTGAA 215

Qy 328 KAIGLVIPELNGKLDAQRPVPTGSVTELVVTLDKNSVDEINAAAMKAASNDSFQYTE 387  
Db 216 KAIGLVIPELNGKLDAQRPVPTGSVTELVAVLEKETSVEEINAAAMKAANDSYGTYE 275

Qy 388 DPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLYKVSWYDNEMSYTAQLVRTLEYFAKIA 447  
Db 276 DPIVSSDIVGMAYGSLFDATQTKVQTVDGNQLVKVSVWYDNEMSYTAQLVRTLEYFAKIA 335

Qy 448 K 448  
Db 336 K 336

RESULT 13  
US-10-134-297-8

; Sequence 8, Application US/10134297  
; Publication No. US20030165524A1  
; GENERAL INFORMATION:  
; APPLICANT: Bolton, Alexandra J.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael A.  
; APPLICANT: Potter, Andrew A.  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GpC PROTEIN AGAINST  
; STREPTOCOCCUS INFECTION  
; FILE REFERENCE: 9000-0055.20  
; CURRENT APPLICATION NUMBER: US/10/134,297  
; CURRENT FILING DATE: 2002-04-26  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Streptococcus uberis

RESULT 12  
US-10-650-369-16  
; Sequence 16, Application US/10650369  
; Publication No. US20040062774A1  
; GENERAL INFORMATION:  
; APPLICANT: Potter, Andrew A.  
; APPLICANT: Perez-Casal, Jose

US-10-134-297-8

Query Match 68.5%; Score 1559.5; DB 14; Length 336;  
Best Local Similarity 73.9%; Pred. No. 2.3e-133;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

Qy 28 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 60

Qy 88 KEGGFEVNGNPIKVSAERDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
Db 61 KDGGEVNGNPIKVSAEKDPE- 81

Qy 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGEIVLE 207  
Db 82 ATGFFAKKAAEKLHANGAKKVITAPGNDVKTUVFNTNHDILDGTETVLSGASCTTN 207  
Db 96 ATGFFAKKAAEKLHANGAKKVITAPGDDDVKTUVFNTNHDILDGTETVLSGASCTTN 155

Qy 208 ATSFPAKKEAAEKLHANGAKKVITAPGNDVKTUVFNTNHDILDGTETVLSGASCTTN 267  
Db 156 CLAPMAKALQDNFGVKQGLMTTIHAYTQDQMLDGPFRGGDLRRARAAGANTIPNSTGAA 215

Qy 268 CLAPMAKALHDAFGIQKGLMTTIHAYTQDQMLDGPFRGGDLRRARAAGANTIPNSTGAA 327  
Db 156 CLAPMAKALQDNFGVKQGLMTTIHAYTQDQMLDGPFRGGDLRRARAAGANTIPNSTGAA 215

Qy 328 KAIGLVIPELINGKLDGAQAQRVPVPTGSVTELVKVSVDEINAAMKAASNSDFGYTE 387  
Db 216 KAIGLVIPELINGKLDGAQAQRVPVPTGSVTELVATLEKDVTVEEVNAAMKAANDSYGYTE 275

Qy 388 DPTVSSDIVGVSYGSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTEYFAKIA 447  
Db 276 DPTVSSDIVGVSYGSLFDATQTKVOTVDGNQLVKVSVWYDNEMSYTSQQLVRLTEYFAKIA 335

Qy 448 K 448  
Db 336 K 336

RESULT 15

US-09-878-781-6

; Sequence 6, Application US/09878781  
; Publication No. US20030082781A1  
; GENERAL INFORMATION:  
; APPLICANT: Bolton, Alexandra J.  
; APPLICANT: Perez-Casal, Jose  
; APPLICANT: Fontaine, Michael  
; APPLICANT: Potter, Andrew A.  
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST  
; FILE REFERENCE: 9000-0055  
; CURRENT APPLICATION NUMBER: US/09/878,781  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 6

LENGTH: 336

TYPE: PRT

ORGANISM: Streptococcus agalactiae

US-09-878-781-6

Query Match 68.4%; Score 1557.5; DB 10; Length 336;  
Best Local Similarity 73.9%; Pred. No. 3.5e-133;  
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

Qy 28 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 87  
Db 1 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRF----- 54

Qy 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
Db 55 -----  
Qy 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGEIVLE 207  
Db 55 -----  
Qy 208 ATSFPAKKEAAEKLHANGAKKVITAPGNDVKTUVFNTNHDILDGTETVLSGASCTTN 267  
Db 96 ATGFFAKKAAEKLHANGAKKVITAPGNDVKTUVFNTNHDILDGTETVLSGASCTTN 155

Qy 268 CLAPMAKALHDAFGIQKGLMTTIHAYTQDQMLDGPFRGGDLRRARAAGANTIPNSTGAA 327  
Db 156 CLAPMAKALQDNFGVKQGLMTTIHAYTQDQMLDGPFRGGDLRRARAAGANTIPNSTGAA 215

Qy 328 KAIGLVIPELINGKLDGAQAQRVPVPTGSVTELVKVSVDEINAAMKAASNSDFGYTE 387  
Db 216 KAIGLVIPELINGKLDGAQAQRVPVPTGSVTELVATLEKDVTVEEVNAAMKAANDSYGYTE 275

QY	388	DPIVSSDIVGSYGSLLFDATOTKVMEVGDGSOLVKVVSWYDNEMSYTAQLVRTLEYFAKIA	447
Db	276	DPIVSSDIVGISYGSLLFDATOTKUQTKVQTDGNQLVKVVSWYDNEMSYTSQQLVRTLEYFAKIA	335
QY	448	K 448	
Db	336	K 336	

Search completed: September 15, 2004, 10:22:15  
Job time : 133 secs

卷之二

Total number of hits satisfying chosen parameters:

Minimum DB seq Length: 0

N; Alternate names: plasmin receptor  
C: Species: *Streptococcus* sp.

C;Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text\_change 33 Jun 2002  
C;Accession: A42963; B42963; JH0750  
R;Lottenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curti  
J. Bacteriol. 174, 5204-5210, 1992  
A;Title: Cloning, sequence analysis, and expression in *Escherichia coli* of a str  
A;Reference number: A42963; MUID:92355491; PMID:1322883  
A;Accession: A42963

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	0.9999999999999998	1	1	1	1	1

Query	Match	Score	DB	Length
Best Local Matches	Match	1652.5;	DB 2;	Length 333;
QY	28	MVVKVGINGFGRIGRLAFRRIQVNVEGVTRINDLTDPPNMLAHLLKYDTTQGRF	QY	87
Db	1	GRIGRLAFRRIQVNVEGVTRINDLTDPPNMLAHLLKYDTTQGRF	Db	60
QY	88	KEGGFEVNGNFIKVSAERDPENIDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKPIKVSAE	QY	147
Db	61	KEGGFEVNGNFIKVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKPIKVSAE	Db	81
QY	148	KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVEIVLE	QY	207
Db	82	KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVEIVLE	Db	95
QY	208	ATSEFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTNHDILDGTETVIVSGASCTTN	QY	267
Db	96	ATSEFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTNHDILDGTETVIVSGASCTTN	Db	155
QY	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAANIVPNSTGAA	QY	327

## ALIGNMENTS

Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGHRRGGDLIRRARAAGAANIVPNSTGAA 215

Qy 328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGYTE 275

RESULT 3

F95235 glyceraldehyde 3-phosphate dehydrogenase [imported] - Streptococcus pneumoniae (strain C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001

C;Accession: F95235

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heinson, J.D.; Umayam, L.A.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, M.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95235

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-335 <KUR>

A;Cross-references: GB:AE005672; PIDN:AAK76079.1; PID:gn14973522; GSPDB:GN00164; TIGR:SP

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP2012

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 67.4%; Score 1535; DB 2; Length 335;

Best Local Similarity 73.6%; Pred. No. 4.4e-93;

Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

Qy 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMNLAHLLKYDTTQGRFDTGTYEV 87

Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLTDPMNLAHLLKYDTTQGRFDTGTYEV 60

Qy 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVKDGFFDVNGKFIKVSAE 147

Db 61 KEGGFEVNGKFIKVSAAERDPPE----- 81

Qy 148 KDPEQIDWATDGVEIVLEIDGTEVKEGGFEVNGQFVVKVSAEREPAENIDWATDGVEIVLE 207

Db 82 ----- 82

Qy 208 ATSFIAKKEAAEKHLHANGAKVVIATAPGGNDVKTFFNTNHDILDGTETVLSGASCTTN 267

Db 96 ATGFFIAKKEAAEKHL-KGGAKKVVITAPGGNDVKTFFNTNHDILDGTETVLSGASCTTN 154

Qy 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGHRRGGDLRRARAAGAANIVPNSTGAA 327

Db 155 CLAPMAKALQDNFGVVEGLMTTIHAYTGDQMLDGHRRGGDLRRARAAGAANIVPNSTGAA 214

Qy 328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGYTE 387

Db 82 ----- 82

Qy 208 ATSFIAKKEAAEKHLHANGAKVVIATAPGGNDVKTFFNTNHDILDGTETVLSGASCTTN 267

Db 96 ATGFFIAKKEAAEKPLHANGAKVVIATAPGGNDVKTFFNTNHDILDGTETVLSGASCTTN 155

Qy 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGHRRGGDLRRARAAGAANIVPNSTGAA 327

Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGHRRGGDLRRARAAGAANIVPNSTGAA 215

Qy 328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGYTE 275

RESULT 4

G98099 glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] - St

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 03-Jun-2002

C;Accession: G98099

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E;e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc; J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.	Qy	28	MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV
A;Reference number: A97872; MUID:21429245; PMID:11544234	Db	1	MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV
A;Accession: G98099			
A;Status: Preliminary			
A;Molecule type: DNA			
A;Residues: 1-359 <KUR>			
A;Cross-references: GB:AE007317; PIDN:AAL00628.1; PID:915459513; GSPDB:GN00174			
C;Genetics:			
A;Gene: gapA			
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase			
C;Keywords: oxidoreductase			
Query Match 67.4%; Score 1535; DB 2; Length 359;			
Best Local Similarity 73.6%; Pred. No. 4.8e-93;			
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;			
28 MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV 87	Qy	208	ATSFQFAKEAAEKHLHANGAKKVITAPGGNDVKTUVFNTNHDILDGTETVISGASCTTN 267
25 MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV 84	Db	96	ATGFFATKEKEQHLHANGAKKVITAPGGSDVKTIVFNTNHEVLDGTETVISGASCTTN 155
268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 327	Qy	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 327
156 CLAPMADTLNKQFGIKVGTMTVHGTYGDQMILDGPHRGCDLRRARAENINVPNSTGAA 215	Db	156	CLAPMADTLNKQFGIKVGTMTVHGTYGDQMILDGPHRGCDLRRARAENINVPNSTGAA 215
328 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSFGYTE 387	Qy	328	KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSFGYTE 387
216 KAIGLVLVPELQGKLGQHAQRVPVPTGSVTELVVTLINKETVDEINAAMKAASNDSFGYNE 275	Db	216	KAIGLVLVPELQGKLGQHAQRVPVPTGSVTELVVTLINKETVDEINAAMKAASNDSFGYNE 275
388 DPIVSSDIVGVSYGSLSFDATOTKVMEVDGSQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447	Qy	388	DPIVSSDIVGVSYGSLSFDATOTKVMEVDGSQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
276 DQIVSSDIVGINSNSLFDATQTEVTSADGAQLVKKVSVWYDNEMSYTSNLVRLTLEYFAKIA 335	Db	276	DQIVSSDIVGINSNSLFDATQTEVTSADGAQLVKKVSVWYDNEMSYTSNLVRLTLEYFAKIA 335
148 KDPEQIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPA	Qy	448	K 448
120 ATGFFAKKEAAEKHL-KGGAKAKVVTAPGGNDVKTUVFNTNHDILDGTETVISGASCTTN 207	Db	120	ATGFFAKKEAAEKHL-KGGAKAKVVTAPGGNDVKTUVFNTNHDILDGTETVISGASCTTN 207
106 -----QIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPA	Db	106	106 -----QIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPA
208 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 267	Qy	208	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 267
120 ATGFFAKKEAAEKHL-KGGAKAKVVTAPGGNDVKTUVFNTNHDILDGTETVISGASCTTN 119	Db	120	120 ATGFFAKKEAAEKHL-KGGAKAKVVTAPGGNDVKTUVFNTNHDILDGTETVISGASCTTN 119
268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 327	Qy	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 327
179 CLAPMAKALQDNFGVVEGLMFTTHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 238	Db	179	CLAPMAKALQDNFGVVEGLMFTTHAYTGDQMILDGPHRGCDLRRARAAGAAINVPNSTGAA 238
328 KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSFGYTE 387	Qy	328	KAIGLVIPELNGKLDGAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSFGYTE 387
239 KAIGLVIPELNGKLDGSAQRVPVPTGSVTELVVTLAVLKNVTDEVNAAMKAASNDSFGYTE 298	Db	239	KAIGLVIPELNGKLDGSAQRVPVPTGSVTELVVTLAVLKNVTDEVNAAMKAASNDSFGYTE 298
388 DPIVSSDIVGVSYGSLSFDATOTKVMEVDGSQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447	Qy	388	DPIVSSDIVGVSYGSLSFDATOTKVMEVDGSQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
299 DPIVSSDIVGMSYGSLSFDATQTKVLDVDGKQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 358	Db	299	DPIVSSDIVGMSYGSLSFDATQTKVLDVDGKQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 358
448 K 448	Qy	448	K 448
359 K 359	Db	359	K 359
RESULT 6			
Q86694 hypothetical protein gapA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)			
C;Species: Lactococcus lactis subsp. lactis			
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001			
C;Accession: G86694			
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Weissenbach, J.; Ehrlich, Genome Res. 11, 731-753, 2001			
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis subsp. Reference number: A86625; MUID:21235186; PMID:113337471			
A;Accession: G86694			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-337 <STO>			
A;Cross-references: GB:AE005176; PID:912723446; PIDN:AAK04657.1; GSPDB:GN00146			
A;Experimental source: strain IL1403			
C;Genetics:			
A;Gene: gapA			
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase			
Query Match 56.2%; Score 1280; DB 2; Length 337;			
Best Local Similarity 61.4%; Pred. No. 2e-76; Mismatches 31; Indels 86; Gaps 2;			
Matches 259; Conservative 31; Gaps 2;			
28 MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV 87	Qy	28	MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV 87
1 MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV 60	Db	1	MVVKVGINGFGRIGRLAERRRIQVNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVEV 60
88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPA	Qy	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPA
61 KEDGFDVNGKFKVTAERNPE-----	Db	61	KEDGFDVNGKFKVTAERNPE-----
148 KDPEQIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPA	Qy	148	KDPEQIDWATDGVEIVLIEDGTVEVKEGGFEVNGQFVKVSAEREPA
82 -----	Db	82	82 -----
208 ATSPFAKKEAAEKHLHANGAKKVITAPGGNDVKTUVFNTNHDILDGTETVISGASCTTN 267	Qy	208	ATSPFAKKEAAEKHLHANGAKKVITAPGGNDVKTUVFNTNHDILDGTETVISGASCTTN 267
96 ATGFFATKEKAEEKHLHPPGGAKKVLITAPGGNDVKTUVFNTNHTILDGTETVISGASCTTN 155	Db	96	ATGFFATKEKAEEKHLHPPGGAKKVLITAPGGNDVKTUVFNTNHTILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDDLRARRAGAANIVPNSTGAA 327  
 Db 156 SLAPMADALNKNFGVKGGTMTTvhSYTGQMTLDPHRRGGDFRRARAVENTVPAASSGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELYVTLDKNVSVDEINAAMKAASNDSEFGYTE 387  
 Db 216 KAIGLVLPPELSELMKGHAQRVSTPTGSITELTVLKEHVTDEINAAMKAASADESFYGVNV 275

QY 388 DPIVSSDIVGVSYGSLSFDATQTKYMEV-DGSQQLVKVYVSWYDNEMSYTAQLVRTLEYFAKI 446  
 Db 276 DEIVSSDIIGMAYGSLFDATLTVTDLKDGGQLVVKTAAYWDNEMSFTAQLIRTLEYFAKI 335

QY 447 AK 448  
 Db 336 AK 337

RESULT 7

S34254 glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Clostridium p  
 C; Species: Clostridium pasteurianum  
 C; Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 03-Jun-2002  
 C; Accession: S34254  
 R; Oster, T.; Assobhei, O.; Scherrer, S.; Branlant, G.; Branlant, C.  
 submitted to the EMBL Data Library, May 1993  
 A; Description: Nucleotide sequence of the glyceraldehyde-3-phosphate dehydrogenase from  
 A; Reference number: S34254  
 A; Accession: S34254  
 A; Molecule type: DNA  
 A; Residues: 1-334 <OST>  
 C; Cross-references: EMBL: X72219; NID: g311923; PIDN: CAA51020.1; PID: g311924  
 C; Superfamily: Glyceraldehyde-3-phosphate dehydrogenase  
 C; Keywords: oxidoreductase

Query Match 54.2%; Score 1234; DB 2; Length 334;  
 Best Local Similarity 59.3%; Pred. No. 2.1e-73;  
 Matches 249; Conservative 34; Mismatches 51; Indels 86; Gaps 3;

QY 29 VVKVGINGFGRIGRLAFRRRIQNVEGVEVTRINDLDPNMLAHLKYDTTQGRFEDGTVEVK 88  
 Db 1 MTRKVALINGFGRIGRLALRRLLEVPGLEWVAINDLTDAKMLAHLFYKYDSSQGRFNGEIEVK 60

QY 89 EGGFEVNGNFIKVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNNGKFKIVSAEK 148  
 Db 61 EGAFVWNGKEVKVFAADPEKL PW-----GEL-----GIDVVLEC 84

QY 149 DPEQIDWATDGVEIVLEIDGTVEVKEGGFITAPGGNDVKTVFNTNHDILDGTEVTLK 208  
 Db 85 -----GEL-----GEL-----GEL-----GIDVVLEC 95

QY 209 TSFFAKBAAEKLHANGAKKVITAPGGFEEVNGQFVKVSAEREPAENIDWATDGVEIVLEA 268  
 Db 96 TGFFFTKKEKAEEAHVRA-GAKKVVVISAPAGNDLKTIVFNVNNEEDLDGTETVIVGASCTTNC 154

QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDDLRARRAGAANIVPNSTGAAK 328  
 Db 155 LAPMAKVLNDKFGIEKGFMTTIAYTNDQNTLDGPHRKGDFFRARAASVAVSIPNSTGAAK 214

QY 329 AIGLVIPELNGKLDGAAQRVPVPTGSVTELYVTLDKNVSVDEINAAMKAASNDSEFGYTE 388  
 Db 215 AIAQVIPELKGKLGDNQAQRVPVPTGSVTELYVTLKVNNTVEEINAAMKEAANESFGYTED 274

QY 389 PISSSDIVGVSYGSLSFDATQTKYMEVDSQQLVKVYVSWYDNEMSYTAQLVRTLEYFAKI 448  
 Db 275 EIVSADVVGISYGSLSFDATLTKIVDVGSQLVKTWSYDNEMSYTSQLVRTLEYFAKI 334

RESULT 8

AC1382 glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria innocua (strain  
 C; Species: Listeria innocua  
 C; Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
 C; Accession: AD1751  
 R; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, U.; Dominguez-Bernal, G.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

A; Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A; Accession: AD1751  
 A; Status: preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-336 <GLA>

A; Cross-references: GB:AL592022; PIDN:CAC97780.1; PID:g16415075; GSPDB:GN00178	Db	61	EGAFVUNGKEYVKAFAEADPERKLPGW-----DL-----	87
C; Genetics:				
A; Gene: gap				
C; Superfamily: glyceraldehyde-3-phosphate dehydrogenase				
Query Match 52.9%; Score 1204.5; DB 2; Length 336;	Db	88		
Best Local Similarity 58.5%; Pred. No. 1.8e-71;	Qy			
Matches 247; Conservative 33; Mismatches 55; Indels 87; Gaps 4;	Db			
28 MVVKVGINGFGRIGRIGRIGRARRIONVEGVEVTRINDLTPNMLAHLJKYDTTQGRFDGTVEV 87	Qy	209	TSEFAKKEAAEKLHLHANGAKKVITAPGGNDVKTVFNTNHDILDGTTETVIGSASCCTNC 268	149 DPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGOFVKVSAEREPAVIDWATDGVEIVLEA 208
1 MTKVGINGFGRIGRIGRIGRARRIONVEGIEVVAINDLTDAKMLAHLJKYDTTQGRFDGEVEV 60	Db	96	TGFFTKKEKZEAHVRA-GAKKVVISAPAGNDLKTIVENVNNEEDLDGTETVIGSASCCTNC 154	GIDVILEC 95
88 KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAA 147	Qy	269	LAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDDLRRARRAGAANIVPNSTGAAK 328	
61 HDGFFNVNGKEVKVLANRNPEELPGW-----DL----- 88	Db	155	LAPMAKVLNDKFGIEKGFMTTIHAFTNDQNTLDPHRKGDLLRARRAAAVSIIIPNSTGAAK 214	
148 KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGVEIVLE 207	Qy	329	AIGLVIPELNGKLDGAAQRVPVPTGSVTELVYTLDKNVSVDEINAAMKAASNDSEFGYTED 388	
89 -----GVDIVLE 95	Db	215	AISQVTPDLAGKLDGNAQRVPVPTGSITELVSVLKRVTEVINAAMKERAADESEFGYTED 274	
208 ATSFPEAKKEAAEKLHLHANGAKKVITAPGGNDVKTVFNTNHDILDGTTETVIGSASCCTN 267	Qy	389	PIVSSDIVGVSYGSLFDATQTKVMEVDSQQLVKKVSVSWYDNEMSYTAQLVRLTLEYFAKIAK 448	
96 CTGFFTAQDKAELHKKA-GAKKVVISAPATGDMKTIIVNVNHETDLGTETVIGSASCCTN 154	Db	275	PIVSADVVGINTYNGSLFDATLTKIVDNGSQLVKTAAWYDNEMSYTSQQLVRLTLEYFAKIAK 334	
RESULT 11				
B82019 glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) C NMA0246 [impo				
C;Species: Neisseria meningitidis				
C;Accession: B82019				
Db: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002				
Db: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002				
Db: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002				
R; Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel				
; Holroyd, S.; Jagels, K.; Leather, S.; Moulle, S.; Mungall, K.; Quail, M.A.; Rajandream,				
Nature 404, 502-506, 2000				
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.				
A;Reference number: A81775; PMID:10761919				
A;Accession: B82019				
A;Status: preliminary				
A;Molecule type: DNA				
A;Residues: 1-334 <PAR>				
A;Cross-references: GB:AL162752; GB:AL157959; NID:97378778; PIDN:CAE83554.1; PID:g737900				
A;Experimental source: serogroup A, strain Z2491				
C;Genetics:				
A;Gene: NMA0246				
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase				
C;Keywords: oxidoreductase				
Query Match 50.3%; Score 1145; DB 2; Length 334;				
Best Local Similarity 56.4%; Pred. No. 1.4e-67;				
Matches 235; Conservative 35; Mismatches 61; Indels 86; Gaps 4;				
28 MVVKVINGFGRIGRIGRARRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87	Qy	28	MVKVINGFGRIGRIGRARRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87	
215 KAIGEVLPPTLKGKLDGAAQRVPVPTGSILTEVTLVKYTVDEVNAAMEAASDPETFGYT 274	Db	1	MSIKVAINGFGRIGRIGRARRIQEKAHGLIEVVAVNDLTPAEMMLHLFKYDSTQGRF----- 54	
387 EDPIVSSDIVGVSYGSLFDATQTKVMEVDSQQLVKKVSVSWYDNEMSYTAQLVRLTLEYFAKII 446	Qy	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147	
275 SDQVVSDDIKGMTFGSLFDETQTKVLTVGDQQLVKTVAWYDNEMSYTAQLVRLTLEYFAKII 334	Db	55	QGTAELKDDAIVVNGKEIKVVFAN 77	
447 AK 448	Qy	148	KDPEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGVEIVLE 207	
335 AK 336	Db	78	-----GEL-----GVDVLE 95	
RESULT 10				
C96987 glyceraldehyde 3-phosphate dehydrogenase, gene gapC [imported] - Clostridium acetobutyli				
C;Species: Clostridium acetobutylicum				
C;Accession: C96987				
R; Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.				
D; Bacteriol. 183, 4823-4838, 2001				
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum				
A;Reference number: A96900; PMID:21359325; PMID:21359325				
A;Accession: C96987				
A;Status: preliminary				
A;Molecule type: DNA				
A;Residues: 1-334 <PAR>				
A;Cross-references: GB:AE001437; PIDN:AAK78686.1; PID:g15023589; GSPDB:GN00168				
A;Experimental source: Clostridium acetobutylicum ATCC824				
C;Genetics:				
A;Gene: CAC0709				
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase				
Query Match 52.8%; Score 1203; DB 2; Length 334;				
Best Local Similarity 57.4%; Pred. No. 2.2e-71;				
Matches 241; Conservative 41; Mismatches 52; Indels 86; Gaps 3;				
29 VVKVGINGFGRIGRIGRARRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEK 88	Qy	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDDLRRARRAGAANIVPNSTGA 327	
1 MAKIAINGFGRIGRIGRARRILEVPGLEVVAINDLTDAKMLAHLFKYDSSQGRFENGTEVK 60	Db	155	CLAPMAAVLQKEFGVVEGLMTTIHAYTGDQNTLDAPHRKGDLLRARRAAALNIVPNSTGA 214	
89 EGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAAK 148	Qy	328	KAIGLVIPELNGKLDGAAQRVPVPTGSVTLDKNVSVDEINAAMKAASNDSEFGYTYE 387	
215 KAIGLVIPELNGKLDGSAQRVPVATGSLTELYSVLERPVTKBEEINAAMKAASESYGYNE 274	Db	215	-----GVDVLE 95	

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Malignani, V.; Pizza, M. Science 287, 1809-1815, 2000  
A; Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R  
A; Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58  
A; Reference number: A81000; MUID:20175755; PMID:10710307

QY	28	MVVKGINGFGRIGRLAFERRI-----QNVVEVTRINDLTDPMNLAHLKYDFTTQGRFDG	Db	98	GFTYSAEKSQAHDLA-GAKKVLIISAPAG-EMKTIIVKVNDDTLGNDTIVVASCTTNCL
Db	1	MTVKIGINGFGRIGRLAFERRIMDLGEETKIDEVVAINDLTTPAMLAHLKYDFTTQGRFDG	QY	270	APMAKALHDAFGI0KGIMTTIAYTGDQMLLDGPHRGDLRRARAGAANTIPNSTGAAKA
QY	84	TVEYKEGGFEVNGNFIKVSAYERDPENIDWATDGVEIVLEALEGTVEVKDGGEDVNGKFIK	Db	156	APMAKALHDSFGIEVGTMTTIAYTGTQSLVDP-RGKDLRASRAAENIIPHTTGAAKA
Db	61	EVSAEDSLVVDGKRYVAEPAQNIWP-----VKN-----	QY	330	IGLVIPELNGKLDGAAQRVPVPTGSVTELVYTLDKNVSVDEINAAMKAA-SNDSFGYTE
QY	144	VSAEKDPEQIDWATDGVEIVLEIDGTVEKEGGFEVNGQFVKVAEREPAWIDWATDGVE	Db	215	IGLVIPELSGRLKGHAQRVPVKTAEVNNALKQATTNNEESFGYTD
Db	93	-----	QY	388	DPIVSSDIVGVSYGSLEDATOTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRLTLEYPAKI
QY	204	IVLEATSSFAKKEAAEKLHANGAKKVVITAPGGNDVWKVVENTNHDILDGTETVSGAS	Db	46	4.6
Db	97	FVLECTGFYTSKAKSQAHLDA-GAKRVLISAPAGNDLKTIVSVNQDTLTADDTIVSAGS	QY	275	EEIVSSDIIIGSHFGSVPDATOTEITAVGDLQLVKTVAWYDNEYGFVTQLIRTLEKFAKL
Db	9	-----	Db	333	3.33
QY	264	CTTNCLAPMAKALHDAFGI0KGIMTTIAYTGDQMLLDGPHRGDLRRARAGAANTIPNS	Db	Search completed: September 15, 2004, 10:27:49	
Db	156	CTTNCLAPMANALNKEFGIQVGTMTTIIAYTATQKVLDGPDRGNNFRNARAEEAENIIPHS	Job time : 43 secs		
QY	324	TGAAKAIGLVLPELINGKLDGAAQRVPVPTGSVTELVYTLDKNVSVDEINAAMKAAASNDSF	Db		
Db	216	TGAAKAIGLVLPELINGKLDGHAQRVPVKGDSSETELVYTLDKKVTAEEVNAAMKKYESPSF	QY	383	3.83
QY	384	GYTEDPIVSSDIVGVSYGSLEDATOTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRLTLEYF	Db	275	4.43
Db	276	AYNADQIVSTDVLMGTTAGSIFDPTQTQVITAGDKQLVWKTVAWYDNEYSTFCQMVRTLLHF	QY	335	3.35
QY	444	AKI 446	Db	336	ATL 338
Db	3	-----	Db	54	-----
RESULT 15					
F90881	glyceraldehyde 3-phosphate dehydrogenase C [imported] - Escherichia coli (strain O157:H7)				
C;Species: Escherichia coli					
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001					
C;Accession: F90881					
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.					
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.					
DNA Res. 8, 11-22, 2001					
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene					
A;Reference number: A99629; MUID:21156231; PMID:11258796					
A;Accession: F90881					
A;Status: preliminary					
A;Molecule type: DNA					
C;Residues: 1-333 <HAY>					
A;Cross-references: GB:BA000007; PIDN:BAB35445.1; PID:913361488; GSPDB:GN00154					
A;Experimental source: strain O157:H7, substrate RIMD 0509952					
C;Genetics:					
A;Gene: ECs202					
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase					
Query Match 41.3%; Score 941.5; DB 2; Length 333;					
Best Local Similarity 47.5%; Pred. No. 2.7e-54; Mismatches 53; Indels 91; Gaps 7;					
Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;					
31 KVGINGFGRIGRLAFERRIONVE-GVEVTRINDLTDPMNLAHLKYDFTTQGRFDGTVEVKE					
3 KVGINGFGRIGRLVRLLEVKSNNIDVVAINDLTSKPLAYLLKHDNSYGP-----					
90 GGFEVNGNFIKVSAYERDPENIDWATDGVEIVLEALEGTVEVKDGGEDVNGKFIKVAEKD					
Db	55	-----	Db	54	-----
QY	150	PEQIDWATDGVEIVLEIDGTVEVKDGGEDVNGKFIKVAEKD	Db	209	20.9
Db	55	-----	Db	54	-----
QY	210	SFFAKKEAAEKLHANGAKKVVITAPGGNDVWKVVENTNHDILDGTETVSGASCTTNCL	Db	97	2.69

This Page Blank (uspto)

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:11:48 ; Search time 25 Seconds  
 (without alignments)  
 933.097 Million cell updates/sec

Title: US-10-650-369-22  
 Perfect score: 2278  
 Sequence: 1 MKKTTGILLLAVILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing First 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1650.5	72.5	335	1 G3P_STRPY	P50467 streptococc
2	1646.5	72.3	335	1 G3P_STRP3	Q8k8m9 streptococc
3	1553.5	68.2	335	1 G3P_STREQ	Q59906 streptococc
4	1280	56.2	337	1 G3P_LACLA	P52987 lactococcus
5	1234	54.2	334	1 G3P_CLOPA	Q59309 clostridium
6	1203	52.8	334	1 G3P_CLOAB	052631 clostridium
7	1138.5	50.0	336	1 G3P1_STAEP	Q8cpv5 staphylococ
8	1132.5	49.7	336	1 G3P1_STAAM	Q9z5c5 staphylococ
9	955.5	41.9	338	1 G3P_LACDE	032755 lactobacill
10	941.5	41.3	333	1 G3P3_ECO57	P58072 escherichia
11	940.5	41.3	333	1 G3P3_ECOLI	P33898 escherichia
12	870.5	38.2	337	1 G3P_MYCPN	P75358 mycoplasma
13	854.5	37.5	337	1 G3P_MYCGE	P47543 mycoplasma
14	850	37.3	334	1 G3P_CORGL	Q01651 corynebacte
15	844	37.1	334	1 G3P1_BACSU	P09124 bacillus su
16	840	36.9	334	1 G3P_BACME	P23722 bacillus me
17	825	36.2	334	1 G3P_BACST	P00362 bacillus st
18	820.5	36.0	335	1 G3P_BORBU	P46795 borrelia bu
19	818.5	35.9	332	1 G3P_THEMEA	P17721 thermotoga
20	793.5	34.8	342	1 G3P_AQUAE	067161 aquifex aeo
21	790	34.7	336	1 G3P_STRCO	Q9z518 streptomyce
22	786.5	34.5	339	1 G3P_MYCTU	006822 mycobacteri
23	783.5	34.4	339	1 G3P_MYCLE	P46713 mycobacteri
24	777.5	34.1	339	1 G3P_MYCAV	P94915 mycobacteri
25	763.5	33.5	330	1 G3P1_SALTY	P24165 salmonella
26	759.5	33.3	330	1 G3P1_ECOLI	P06977 escherichia
27	754	33.1	339	1 G3P_HAEIN	P44304 haemophilus
28	743.5	32.6	330	1 G3P_TRYBB	P10097 trypanosoma
29	74.0	32.5	332	1 G3P_RALSO	P52694 ralstonia s
30	73.9	32.4	332	1 G3P_STRAU	Q59800 streptomyce
31	73.7	32.4	337	1 G3P1_RHIRA	Q9c136 rhizomucor
32	73.7	32.4	337	1 G3P_MONAN	P53430 monascus an
33	73.4	32.2	336	1 G3P1_SCHPO	P78958 schizosacch

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
 CC -!- PATHWAY: Second phase of glycolysis; first step.  
 CC -!- SUBUNIT: Homotetramer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cyttoplasmic.  
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; M95569; AAA26953.1; -. DR EMBL; AE006494; AAK33348.1; -. DR HSSP; P00362; 1GDI.  
 DR InterPro; IPR000173; GAP\_dhydrogenase.  
 DR InterPro; IPR006424; GAPDH-I.  
 DR Pfam; PF00044; gpdh; 1.  
 DR Pfam; PF02800; gpdh\_C; 1.  
 DR PRINTS; PR00078; G3PDHDRGNASE.  
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
 DR PROSITE; PS00071; GAPDH; 1.  
 DR GLYCOLYSIS; Oxidoreductase; NAD; Complete proteome.

FT INIT MET 0 0  
 FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).  
 FT ACT\_SITE 178 ACTIVATES THIOL GROUP DURING CATALYSIS  
 FT CONFLICT 261 261 (BY SIMILARITY).  
 SQ SEQUENCE 335 AA; 35811 MW; F06006EE253C8A3F CRC64;

Query Match 72.5%; Score 1650.5; DB 1; Length 335;  
 Best Local Similarity 79.3%; Pred. No. 1.2e-93;  
 Matches 333; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

QY 29 VVKVGLNGFGRIGRLAFLRRIQVNRVTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 88  
 DB 1 VVKVGLNGFGRIGRLAFLRRIQVNRVTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 60  
 QY 89 EGGFEVNGNFIKVSAERDPENIDWATDGVEIVLEALEGTVKQDGGFDRVNGKFKVSAEK 148  
 DB 61 EGGFEVNGNFIKVSAERDPE----- 80  
 QY 149 DPEQIDWATDGVEIVLEIDGGTVEVKGGFREVNGQFVKVSAEREPAVIDWATDGVEIVLEA 208  
 DB 81 -----NIDWATDGVEIVLEA 95

QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGGNDVVKTVVENTNHDILDGTETVIGSACTN 268  
 DB 96 TGFFAKKEAAEKHLHANGAKKVVITAPGGNDVVKTVVENTNHDILDGTETVIGSACTN 155  
 QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAANTVPNSTGAAK 328  
 DB 156 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAANTVPNSTGAAK 215  
 QY 329 AIGLVLPELNGKLDDGAQRVFPVPTGSVTELVVTLDKNNVSDEINAAMKAASNSDFGTYED 388  
 DB 216 AIGLVLPELNGKLDDGAQRVFPVPTGSVTELVVTLDKNNVSDEINAAMKAASNSDFGTYED 275  
 QY 389 PIVSSDIVGVSYGSSLFDATQTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIAK 448  
 DB 276 PIVSSDIVGVSYGSSLFDATQTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335

CC DT 28-FEB-2003 (Rel. 41; Created)  
 CC DT 28-FEB-2003 (Rel. 41; Last sequence update)  
 CC DT 15-MAR-2004 (Rel. 43; Last annotation update)  
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)  
 DE (Plasminogen-binding protein) (Plasmin receptor).  
 GN GAP OR PLR OR SPYMS\_0201 OR SPS0207.  
 OS Streptococcus pyogenes (serotype M3).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OC NCBI\_TaxID=198466;  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MGAS315 / Serotype M3;  
 RX MEDLINE=22133808; PubMed=12122206;  
 RA Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S., Parkins L.D., Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D., Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M., Schlievert P.M., Musser J.M.; RT "Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).  
 RL RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SSI-1 / Serotype M3;  
 RX MEDLINE=22683278; PubMed=12799345;  
 RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y., Okahashi N., Kawabata S., Shiba T., Yasunaga T., Hayashi H., Hattori M., Hamada S.; RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution.";  
 RT Genome Res. 13:1042-1055 (2003).  
 RL CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN (BY similarity).  
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
 CC -!- PATHWAY: Second phase of glycolysis; first step.  
 CC -!- SUBUNIT: Homotetramer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cyttoplasmic.  
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AE014140; IAM78808.1; -. DR EMBL; AP005141; BAC63302.1; -. DR InterPro; IPR000173; GAP\_dhydrogenase.  
 DR IPR006424; GAPDH-I.  
 DR Pfam; PF00044; gpdh; 1.  
 DR PRINTS; PR00078; G3PDHDRGNASE.  
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
 DR PROSITE; PS00071; GAPDH; 1.  
 DR GLYCOLYSIS; Oxidoreductase; NAD; Complete proteome.  
 KW BY SIMILARITY.  
 FT INIT MET 0 0 BY SIMILARITY.  
 FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).  
 FT ACT\_SITE 178 ACTIVATES THIOL GROUP DURING CATALYSIS  
 FT CONFLICT 261 261 (BY SIMILARITY).  
 SQ SEQUENCE 335 AA; 35811 MW; F06006EE253C8A3F CRC64;

Query Match 72.3%; Score 1646.5; DB 1; Length 335;  
 Best Local Similarity 79.0%; Pred. No. 2.1e-93;  
 Matches 332; Conservative 1; Mismatches 2; Indels 85; Gaps 1;

QY 29 VVKVGLNGFGRIGRLAFLRRIQVNRVTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 88

RESULT 2  
 G3P STRP3 STANDARD;  
 ID G3P STRP3  
 AC Q8K8M9;

Db	1	VVRVGGINGFGRIGRLAFRRRIQNIEGVETRINDLDPNMLAHLKKYDTTQGRFDTGTVVK	60	DR InterPro; IPR000173; GAP dehydrogenase.
QY	89	EGGFEVNGNFIKVSAAERDPENIDWATDGVIEVLEALEGTVEKDGGFDVNGKFIFKVSAAEK	148	DR InterPro; IPR006424; GAPDH-1.
Db	61	EGGFEVNGNFIKVSAAERDPE-----	80	DR Pfam; PF00044; gpdh; 1.
Db	149	DPEQIDWATDGVIEVLEIDGTVEVKKEGVQFVVKVSAAEREPANIDWATDGVIEVLEA	208	DR Pfam; PF02800; gpdh; C; 1.
QY	81	-	-	DR PRINTS; PRO0078; G3PDHDRGNASE.
Db	209	TSEFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVIGASCTTNC	268	DR TIGRFAMS; TIGR01534; GAPDH-1; 1.
QY	96	TGFFAKKEAAEKHLHNTGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVIGASCTTNC	155	DR PROSITE; PS00071; GAPDH; 1.
Db	269	LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK	328	KW GLYCOLYSIS; Oxidoreductase; NAD. BY SIMILARITY.
QY	156	LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK	215	FT INIT MET 0 0 GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
Db	329	AIGLVIPELNGKLDGAAQRVPVPTGSVTELVTLDKNVSVDEINAAMKAASNDSGFYTED	388	FT BINDING 151 151 ACTIVATES THIOL GROUP DURING CATALYSIS
QY	216	AIGLVIPELNGKLDGAAQRVPVPTGSVTELVTLDKNVSVDEINAAMKAASNDSGFYTED	275	FT ACT_SITE 178 178 (BY SIMILARITY).
Db	389	PIVSSDIVGVSYGSLFDATQTKMVEVDGSQQLVKVVSWDNEMSYTAQLVRTLEYFAKIAK	448	FT SEQ 335 AA; 35739 MW; FE7ACFDFD7663E46 CRC64;
QY	276	PIVSSDIVGVSYGSLFDATQTKMVEVDGSQQLVKVVSWDNEMSYTAQLVRTLEYFAKIAK	335	Query Match 68.2%; Score 1553.5%; DB 1; Length 335; Best Local Similarity 76.0%; Pred. No. 9.5e-88; 2; Mismatches 14; Indels 85; Gaps 1;
Db	81	RESULT 3	335 AA.	Matches 319; Conservative 319; NIDWATDGVIEVLEA 95
QY	81	G3P_STREQ	STANDARD; PRT; 335 AA.	Qy 29 VVKVGINGFGRIGRLAFRRRIQNVGEGEVTRINDLDPNMLAHLKKYDTTQGRFDTGTVVK 88
Db	81	ID G3P_STREQ	STANDARD; PRT; 335 AA.	Qy 1 VVKVGINGFGRIGRLAFRRRIQNVGEGEVTRINDLDPNMLAHLKKYDTTQGRFDTGTVVK 60
QY	81	AC Q59506;	STANDARD; PRT; 335 AA.	Db 61 EGGFEVNGNFIKVSAAERDPE----- 80
DT	01-NOV-1997	(Rel. 35, Created)		Qy 149 DPEQIDWATDGVIEVLEIDGTVEVKKEGVQFVVKVSAAEREPANIDWATDGVIEVLEA 208
DT	01-NOV-1997	(Rel. 35, Last sequence update)		Db 81 ----- NIDWATDGVIEVLEA 95
DT	28-FEB-2003	(Rel. 41, Last annotation update)		Qy 209 TSSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVIGASCTTNC 268
DE		Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)		Db 96 TGGFFAKKEAAEKPLHANGAKKVVITAPGGNDVKQLFSTLTSILDGTETVIGASCTTNC 155
DE		(Plasminogen-binding protein) (Plasmin receptor).		Qy 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK 328
GN		GAP OR GAPC.		Db 156 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGARK 215
OS		Streptococcus equisimilis.		Qy 329 AIGLVIPELNGKLDGAAQRVPVPTGSVTELVTLDKNVSVDEINAAMKAASNDSGFYTED 388
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				Db 216 AIGLVIPELNGKLDGAAQRVPVPTGSVTELVTLDKNVSVDEINAAMKAASNDSGFYTED 275
OC		Streptococcus.		Qy 389 PIVSSDIVGVSYGSLFDATQTKMVEVDGSQQLVKVVSWDNEMSYTAQLVRTLEYFAKIAK 448
OX		NCBI_TAXID=119602;		Db 276 PIVSSDIVGVSYGSLFDATQTKMVEVDGSQQLVKVVSWDNEMSYTAQLVRTLEYFAKIAK 335
RN		[1]		[1] SEQUENCE FROM N.A.
RP		SEQUENCE FROM N.A.		RP STRAIN=LMO230;
RC		STRAIN=96305364; PubMed=8706717;		RC MEDLINE=95291425; PubMed=7773380;
RX		Gase K., Gase A., Schirmer H., Malke H.;		RA Cancilla M.R., Hillier A.J., Davidson B.E.;
RA		"Cloning, sequencing and functional overexpression of the		RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene, gap: further evidence for strongly biased codon usage in glycolytic pathway genes.";
RT		Streptococcus equisimilis H46A gapC gene encoding a		RT GAP OR LL0559.
RT		glyceraldehyde-3-phosphate dehydrogenase that also functions as a		OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
RT		plasminogen-binding protein. Purification and biochemical		OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
RT		characterization of the protein. ";		OC NCBI_TaxID=1360;
RT		Eur. J. Biochem. 239:42-51(1996).		RN [1]
RT		-!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.		RP SEQUENCE FROM N.A.
CC		-!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +		RC MEDLINE=95291425; PubMed=7773380;
CC		NAD (+) = 3-phospho-D-glyceroyl phosphate + NADH.		RA Cancilla M.R., Hillier A.J., Davidson B.E.;
CC		-!- PATHWAY: Second phase of glycolysis; First step.		RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene, gap: further evidence for strongly biased codon usage in glycolytic pathway genes.";
CC		-!- SUBUNIT: Homotetramer (By similarity).		RT Microbiology 141:1027-1036 (1995).
CC		-!- SUBCELLULAR LOCATION: Cytoplasmic.		RL [2]
CC		-!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate		RN SEQUENCE FROM N.A.
CC		dehydrogenase family.		RC MEDLINE=21235186; PubMed=11337471;
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration		RC
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -		RC
CC		the European Bioinformatics Institute. There are no restrictions on its		RA
CC		use by non-profit institutions as long as its content is in no way		RT
CC		modified and this statement is not removed. Usage by and for commercial		RT
CC		entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		RT
CC		or send an email to license@isb-sib.ch.)		RL
CC		DR EMBL; X97788; CAA66377; 1.		RN
DR		DR EMBL; Y12602; CAA73174; 1.		RP
DR		DR PIR; S71350; S71350.		RC
DR		DR HSSP; P00362; 1G1.		RX

Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K., Weissenbach J., Ehrlich S.D., Sorokin A.; "The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403."; *Genome Res.* 11:731-753 (2001).

-!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

-!- PATHWAY: Second phase of glycolysis; first step.

-!- SUBUNIT: Homotetramer (By similarity).

-!- SUBCELLULAR LOCATION: Cytoplasmic.

-!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce@isb-sib.ch>) or send an email to license@isb-sib.ch.

2008 ATSEFAKKEAAFEKH HANCAKKIUTANCONDUITY DCEPTEYUOOGGEE

EDITIONES VITAE IN GASCEN 267

96 ATGFFATKEKAEEKHLLPGGGAKKVLITAPGGNDVKTVVFNNTNHTILDGTETVISAGSCTTN 155

268 CLAPMAKALHDAFGIQQKGLMTTIHAYTGDOMILDGPHRGCGDLRRARAGAANTVPNSTGAA 327

EE 514 ADAM DAY 10/10/2013

138 SHAFNALNNF GVRGGMIVVHSY1GDBQMTLDGCPHRGGDFRRARAAAENIVPASSGAA 215

328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEGYTE 387

216 KAIGLVPESLGIIMKGHADORYSTPTGSITELVITVIEKHNUTYDETEWYDADDCEGAWI 225

Z/5 ENGINEERING INSTITUTE OF THE UNIVERSITY OF TORONTO

88 DPIVSSDIVGVSYGSLEFDATQTKVMEV-DGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKI 446

76 DEIVSSDIIIGMAYGSLFDATLTVTDLKDGGSOLVKTAAYWDNEMSETAQIIRTTI.EYFAKI 335

RESULT 5	
G3P_CLOPA	STANDARD; PRT; 334 AA.
ID G3P_CLOPA	STANDARD; PRT; 334 AA.
AC Q59309;	
DT 01-NOV-1997 (Rel. 35, Created)	
DT 01-NOV-1997 (Rel. 35, Last sequence update)	
DT 28-FEB-2003 (Rel. 41, Last annotation update)	
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP 17/CP 18).	
DE	
GN GAP	
OS Clostridium pasteurianum.	
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;	
OC Clostridium.	
OX	
RN NCBI_TaxID=1501;	
RN [1]	SEQUENCE FROM N.A.
RA Oster T., Assobhei O., Scherrer S., Branlant G., Branlant C.;	
RL Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.	
RN [2]	
RP PARTIAL SEQUENCE OF 1-26.	
RC STRAIN=W5;	
RX MEDLINE=98291870; PubMed=9629918;	
RA Flengsrød R., Skjeldal L.;	
RT "Two-dimensional gel electrophoresis separation and N-terminal sequence analysis of proteins from Clostridium pasteurianum W5."	
RL Electrophoresis 19:802-806 (1998).	
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.	
CC -!- PATHWAY: Second phase of glycolysis; first step.	
CC -!- SUBCELLULAR LOCATION: Homotetramer (By similarity).	
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate cytoplasmic.	
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.	
CC	
-----	
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <a href="http://www.isb-sib.ch/ans">http://www.isb-sib.ch/ans</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .)	
CC	
DR X72219; CAA51020.1; -.	
DR PIR; S34254; S34254.	
DR HSSP; P00362; 1GD1.	
DR InterPro; IPR000173; GAP_dhydrogenase.	
DR InterPro; IPR006424; GAPDH-I.	
DR Pfam; PF00044; gpdh; 1.	
DR Pfam; PF02800; gpdh_C; 1.	
DR PRINTS; PR00078; G3PDHDRGNASE.	
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.	
DR PROSITE; PS00071; GAPDH; 1.	
DR GLYCOLYSIS; Oxidoreductase; NAD.	
DR BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.	
DR ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.	
DR SEQUENCE 334 AA; 36078 MW; D15905D0DA7F62E7 CRC64;	
DR	
Query Match 54.2%; Score 1234; DB 1; Length 334;	
Best Local Similarity 59.3%; Pred. No. 2.6e-68;	
Matches 249; Conservative 34; Mismatches 51; Indels 86; Gaps 1	
QY 29 VVKVGVINGFGRIGRGLAFRRRIQNVGVEVTRINDLTDPNMLAHILLKYDTTQGRFDGTVEV :	
Db 1 MTKVAVINGFGRIGRGLAARRRILEVPGVVAINDLTDAKMLAHILFKYDSSQGRFNGEIEEV	
QY 89 EGGFEVNGNFVIKVSAAERDPENIDWATDGVEIVLEALEGTVKDGFFDVGNGKFIKVSAE :	
Db 61 EGAFVNGKEVKVFAEADPEKLPW	

QY	149	DPEQIDWATDGVEIVLEIDGTVEKEGGFEVNGQFVKVSAEREPA	NIDWATDGVEIVLEA	208	DR	PIR; C96987; C96987.
DB	85	-	-----GEL-----	95	DR	HSSP; P17721; 1HDG.
QY	209	TSFFAKKEAAEKLHANGAKKVITAPGGNDVKTVFNTNHDILDGTET	EVNGASCTTNC	268	DR	InterPro; IPR000173; GAP dehydrogenase.
DB	96	TGFFTKKEKAEEAHVRA-GAKKVVISAPAGNDLKTIVFVN	NNDLDGTETEVNGASCTTNC	154	DR	InterPro; IPR006424; GAPDH-I.
QY	269	LAPMAKALHDAGFIQKGLMTTHAYTGDQMLDGHRRAGAANTV	PNSTGAAK	328	DR	PFam; PF00044; gpdh; 1.
DB	155	LAPMAKVLDKFGTIEKGFMTTIHAYTNDQNTLDGPHRKGD	FRRARAAVSI	214	DR	PFam; PF02800; gpdh; C; 1.
QY	329	AIGLVIPELNGKLJGAAQRVVPVPTGSVTEVTLJDKNVSYDE	INAMAKAASNDSFGYTED	388	DR	PRINTS; PR00078; G3PDH-RGNASE.
DB	215	AIAQVIPELKGKLQDGNAAQRVVPVPTGSVTEVTLJDKNVSYDE	INAMKEAANESFGYTED	274	DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.
QY	389	PIVSSDIVGVSYGSLFDATQTKYMEVDGSQLYKVWSYDNE	MSYTAQLVRLTLEYFAKIAK	448	DR	PROSITE; PS00071; GAPDH; 1.
DB	275	EIVSADVVGISYGSLFDATLTKIVDVDSQQLVKTVWSYDNE	MSYTSQQLVRLTLEYFAKIAK	334	DR	GLYCOLYSIS; Oxidoreductase; NAD; Complete proteome.
QY	6	G3P_CLOAB	STANDARD;	PRT; 334 AA.	DR	GLYCERALDEHYDE 3-PHOSPHATE.
DB	025631;	AC	052631;	AC	DR	BINDING
DT	16-OCT-2001	(Rel. 40, Created)	DT	16-OCT-2001 (Rel. 40, Last sequence update)	DR	ACT SITE
DT	28-FEB-2003	(Rel. 41, Last annotation update)	DT	28-FEB-2003 (Rel. 41, Last annotation update)	DR	ACT SITE
DE	Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12)	(GAPDH).	DE	Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12)	DR	ACT SITE
GN	GAP OR GAPC OR CAC0709		GN	CAC0709	DR	ACT SITE
OS	Clostridium acetobutylicum.		OS	Clostridium acetobutylicum.	DR	ACT SITE
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;		OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;	DR	ACT SITE
OC	Clostridium.		OC	Clostridium.	DR	ACT SITE
OX	NCBI_TaxID=1488;		OX	NCBI_TaxID=1488;	DR	ACT SITE
RN	[1]	RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.		DR	ACT SITE
RC	STRAIN=ATCC B24 / DSM 792 / VKM B-1787;		RC	STRAIN=ATCC B24 / DSM 792 / VKM B-1787;	DR	ACT SITE
RX	MEDLINE=21359325; PubMed=11466286;		RX	MEDLINE=99392446; PubMed=10463150;	DR	ACT SITE
RA	Noelling J., Bretton M.V., Omelchenko M.V., Makarova K.S., Zeng Q.,		RA	Noelling J., Bretton M.V., Omelchenko M.V., Qiu D., Hitti J., Wolf Y.I., Daly M.J.,	DR	ACT SITE
RA	Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I., Daly M.J.,		RA	Tatusov R.L., Sabathé F., Doucette-Stamm L., Soucaille P., Daly M.J.,	DR	ACT SITE
RA	Tatusov R.L., Sabathé F., Doucette-Stamm L., Soucaille P., Daly M.J.,		RA	Bennett G.N., Koonin E.V., Smith D.R.,	DR	ACT SITE
RT	"Genome sequence and comparative analysis of the solvent-producing		RT	"Genome sequence and comparative analysis of the solvent-producing	DR	ACT SITE
RT	bacterium Clostridium acetobutylicum."		RT	bacterium Clostridium acetobutylicum."	DR	ACT SITE
RT	J. Bacteriol. 183:4823-4838(2001).		RT	J. Bacteriol. 183:4823-4838(2001).	DR	ACT SITE
RL	-!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +		RL	-!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +	DR	ACT SITE
CC	NAD (+) = 3-phospho-D-glyceroyl phosphate + NADH.		CC	NAD (+) = 3-phospho-D-glyceroyl phosphate + NADH.	DR	ACT SITE
CC	-!- PATHWAY: Second phase of glycolysis; first step.		CC	-!- PATHWAY: Second phase of glycolysis; first step.	DR	ACT SITE
CC	-!- SUBUNIT: Homotetramer.		CC	-!- SUBUNIT: Homotetramer.	DR	ACT SITE
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).		CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).	DR	ACT SITE
CC	-!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate		CC	-!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate	DR	ACT SITE
CC	dehydrogenase family.		CC	dehydrogenase family.	DR	ACT SITE
CC	-----		CC	-----	DR	ACT SITE
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	DR	ACT SITE
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	DR	ACT SITE
CC	the European Bioinformatics Institute. There are no restrictions on its		CC	the European Bioinformatics Institute. There are no restrictions on its	DR	ACT SITE
CC	use by non-profit institutions as long as its content is in no way		CC	use by non-profit institutions as long as its content is in no way	DR	ACT SITE
CC	modified and this statement is not removed. Usage by and for commercial		CC	modified and this statement is not removed. Usage by and for commercial	DR	ACT SITE
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	DR	ACT SITE
CC	NAD (+) = 3-phospho-D-glyceroyl phosphate + NADH.		CC	NAD (+) = 3-phospho-D-glyceroyl phosphate + NADH.	DR	ACT SITE
CC	-!- PATHWAY: Second phase of glycolysis; first step.		CC	-!- PATHWAY: Second phase of glycolysis; first step.	DR	ACT SITE
CC	-!- SUBUNIT: Homotetramer (By similarity).		CC	-!- SUBUNIT: Homotetramer (By similarity).	DR	ACT SITE

RESULT 7  
G3P1\_STAEP STANDARD; PRT; 336 AA.

1D G3P1\_STAEP STANDARD; PRT; 336 AA.

AC Q8CPY5; 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DT 15-MAR-2004 (Rel

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
 CC dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 use by non-profit institutions and the European Bioinformatics Institute. There are no restrictions on its  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

DR EMBL; AE016745; AAO04154.1; -.  
 DR InterPro; IPR006424; GAPDH-I.  
 DR InterPro; IPR000173; GAP\_dhdrogenase.  
 DR Pfam; PF00044; Gpdh; 1.  
 DR Pfam; PF02800; Gpdh\_C; 1.  
 DR PRINTS; PR00078; G3PDHDRGNASE.  
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
 DR PROSITE; PS00071; GAPDH; 1.  
 KW Glycolysis; Oxidoreductase; NAD; Complete proteome.  
 FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY  
 SIMILARITY).  
 FT ACT\_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS  
 (BY SIMILARITY).  
 FT SEQUENCE 336 AA; 36190 MW; A962202D02AB5767 CRC64;  
 SQ

Query Match 50.0%; Score 1138.5; DB 1; Length 336;  
 Best Local Similarity 55.7%; Pred. No. 1.7e-62;  
 Matches 235; Conservative 39; Mismatches 61; Indels 87; Gaps 4;.

QY 28 MVVKVGINGFGRIGRLAARRIONVEGVEVTRINDLTDPNMLAHLLKYDTTQGRFDGTVEV 87  
 1 MAIKVAINGFGRIGRLAARRIQTDEPAGKLPW----GDLDD---- 89  
 QY 88 KEGGFEYNGNFIKVSAAERDPENIDWATDGVETIVALEGTVVKDGGFDVNGKFLIKVSAB 147  
 61 TEGGFRVNGKEIKSFDEPDAGKLPW----GDLDD---- 89  
 QY 148 KDPEQIDWATDGVETIVLEIDGTVEVKEGGFEVNGQFVKVSABREPANIDWATDGVEIVILE  
 90 ---- IDVYLE 95  
 DB 208 ATSEFAKKEAAEKHLHANGAKKVVITAPGGNDVFKTUVFNTNHDILDGTETVISGASCTTN 267  
 DB 96 CTGFYTDKEKAQAHIDA-GAKKVVLISAPAKGDVKTIVFNTNHDILDGETUVSGASCTTN 154  
 QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRARAAGANTIPNSTGAA 327  
 DB 155 SLAPVAKVLSDEFGLVGEFMTTIHAYTGDQNTQDAPHRKGDKERARAEEENIIPNSTGAA 214  
 QY 328 KAIGLVIPELNGKLDGAQRVYPVPTGSVTELVYVTLDK-NVSVDENAAAMKAASNDSFGYI 386  
 DB 215 KAIGKVIPEIDGKLDGGAQRVYPVATGSLTEITVYVLDKQDVTVEQUNSAMKQASDESFGYT 274  
 QY 387 EDPIVVSDIVGVSYGSLFDATOTKVMEDSQLVYKVVSWYDNEMSYTAQLVRTLEYFAKI 446  
 DB 275 EDEIVVSDIVGMYGSLFDATQTRVMTVGDRQLVKVAAWYDNEMSYTAQLVRTLHLAEL 334  
 QY 447 AK 448  
 DB 335 SK 336

RESULT 8  
 G3P1\_STAAM STANDARD; PRT; 336 AA.  
 ID G3P1\_STAAM  
 AC Q9Z5C5;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1).  
 GN GAP OR GAP OR SA0772 OR SA0772 OR MW0734;

OS Staphylococcus aureus (strain Mu50 / ATCC 700699),  
 OS Staphylococcus aureus (strain N315),  
 OS Staphylococcus aureus (strain MW2), and  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=158878, 158879, 196620, 1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RN STRAIN=Mu50 / ATCC 700699, and N315;  
 RC MEDLINE=21311952; PubMed=11418146;  
 RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,  
 RT "Whole genome sequencing of meticillin-resistant Staphylococcus  
 aureus." ; Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21311952; PubMed=11418146;  
 RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,  
 RT "Whole genome sequencing of meticillin-resistant Staphylococcus  
 aureus." ; Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RL [3]  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=22040717; PubMed=12044378;  
 RX Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
 RA Yamamoto K., Hiramatsu K.;  
 RT "Genome and virulence determinants of high virulence community-  
 acquired MRSA." ; Lancet 359:1819-1827 (2002).  
 RL [4]  
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
 CC -!- PATHWAY: Second phase of glyccolysis; first step.  
 CC -!- SUBUNIT: Homotetramer (BY similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (BY similarity).  
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
 CC dehydrogenase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; A0J13320; CAB38645.1; -.  
 DR EMBL; AP003360; BAB56934.1; -.  
 DR EMBL; AP003131; BAB41960.1; -.  
 DR EMBL; AP004824; BAB94599.1; -.  
 DR PIR; E89850; E89850.  
 DR SWISS-2DPAGE; Q9Z5C5; STAAN.  
 DR HSSP; P17721; 1HDG.  
 DR InterPro; IPR006424; GAPDH-I.  
 DR InterPro; IPR000173; GAP\_dhdrogenase.  
 DR PROSITE; PS00071; GAPDH; 1.  
 DR GLYCOLYSIS; Oxidoreductase; NAD; Complete proteome.  
 DR GLYCERALDEHYDE 3-PHOSPHATE (BY  
 DR SIMILARITY).  
 DR ACTIVATES THIOL GROUP DURING CATALYSIS  
 DR ACT\_SITE 178 178  
 DR SQ SEQUENCE 336 AA; 36281 MW; 37A6CEA9376779E5 CRC64;

Query Match	Score	Length	DB	1	CC
Best Local Similarity	49.7%	336;	Pred.	No. 3.9e-62;	
Matches 231; Conservative	54.7%	44;	Mismatches	60;	Indels
					87; Gaps 4;
28 MVVKVINGFGRIGRILAFRRRIQNVGVEVTRINDLTDPNMLAHLLKYDTTQGRFDGTVEV	87				
1 MAVRVAINGFGRIGRILAFRRRIQEVGLEVAVNDLDDMLAHLLKYDTMQGRFTGEVEV	60				
88 KEGGFEVNGNFIKVSAERDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE	147				
61 VDGGFRVNGKREVKSFSEPDASKLPWK-----DIN-----DIN-----DIN-----DIN-----	89				
148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGVEIVLE	207				
					207
90 -----IDVVLE 95					
208 ATSSFFAKKEAEKHLHANGAKKVITAPGGNDVVKTVFVNTNHDILDGTETVSGASCTTN	267				
96 CTGFYTDKDKDQAHIETGAKKVLISAPATGDLKTIVFNTNHQELDGSETVSGASCTTN	154				
268 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMLDGPHRGDDLRARRAGAANIVPNSTGAA	327				
155 SLAPVAKVLNDDDFGLVEGLMTTIIHAYTGDQNTQDAPHRKGDKRRARAENIIPNSTGAA	214				
328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDK-NVSVIDEINAAMKAASNDSFGYT	386				
215 KAIGKVIPEIDGKLDGGAQRVPVATGSLTTELTVLEKQDVTVEQVNNEAMKNASNESFGYT	274				
387 EDPPIVSSDIVGVSYGSLSFDATQTKVMEVDGSQLVKTWSWYDNEMSYTAQLVRTLEYFAKI	446				
275 EDEIVSSDVVGMTYGSLSFDATQTRVMSVGDRLQLVKVAAWYDNEMSYTAQLVRTLAYLAEI	334				
					334
447 AK 448					
335 SK 336					
RESULT 9					
D_G3P_LACDE	STANDARD;	PRT;	338 AA.		
C_O32755;					
T_16-OCT-2001	(Rel. 40, Created)				
T_16-OCT-2001	(Rel. 40, Last sequence update)				
T_28-FEB-2003	(Rel. 41, Last annotation update)				
T_Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)					
GAP.					
Lactobacillus delbrueckii (subsp. bulgaricus).					
Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;					
Lactobacillus.					
NCBI_TaxID=1585;					
[1]					
SEQUENCE FROM N.A.					
STRAIN=B107;					
MEDLINE=98240227; PubMed=9579064;					
A Branny P., Delatorre F., Garel J.R.;					
T "An operon encoding three glycolytic enzymes in Lactobacillus delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase.";					
T Microbiology 144:905-914 (1998).					
T !- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD (+) = 3-phospho-D-glyceroyl phosphate + NADH.					
T !- PATHWAY: Second phase of glycolysis; first step.					
T !- SUBUNIT: Homotetramer (By similarity).					
T !- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).					
T !- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.					
RESULT 10					
G3P3_ECO57	STANDARD;	PRT;	333 AA.		
ID_G3P3_ECO57					
AC_P58072;					
DT_16-OCT-2001 (Rel. 40, Created)					
DT_16-OCT-2001 (Rel. 40, Last sequence update)					
DT_28-FEB-2003 (Rel. 41, Last annotation update)					
DE_Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).					
GN_GAPC OR Z2304 OR ECS2022.					
OS_Escherichia coli O157:H7.					
OC_Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;					
OC_Enterobacteriaceae; Escherichia.					
OC_Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;					
OC_Enterobacteriaceae; Escherichia.					
NCBI_TaxID=83334;					
[1]					
SEQUENCE FROM N.A.					
RP					

RC	STRAIN=O157:H7 / ATCC 700927;	Db	98 GFYTSAAEKSQAHDLA-GAKVVLISAPAG-EMKTTIVYKVNDTIVSWASCTTNCL 155
RX	MEDLINE=21074935; PubMed=11206551;	RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,	RA	Rose D.J., Mayhew G.F., Gregor J., Kirkpatrick H.A.,
RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,	RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,	RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,	RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA	Welch R.A., Blattner F.R.;	RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL	Nature 409:529-533 (2001).	RL	Nature 409:529-533 (2001).
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=O157:H7 / RIMD 0509952;	RC	STRAIN=21156231; PubMed=11258796;
RX	MEDLINE=21156231; PubMed=11258796;	RA	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tobe T.,	RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA	Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,	RA	Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.,	RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.,
RT	"Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";	RT	O157:H7 and genomic comparison with a laboratory strain K-12.;"
RL	DNA Res. 8:11-22 (2001).	CC	-!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC	-!- PATHWAY: Second phase of glycolysis; first step.	CC	-!- SUBUNIT: Homotetramer (By similarity).
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).	CC	-!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.
CC	-!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
CC	-	DR	EMBL; AE005364; AAG56359.1; -.
CC	-	DR	EMBL; AP002557; BAB35445.1; -.
CC	-	DR	PIR; C85737; C85737.
CC	-	DR	PIR; F90881; F90881.
CC	-	DR	HSSP; P17721; 1HDG.
CC	-	DR	InterPro; IPR00173; GAP dhydrogenase.
CC	-	DR	InterPro; IPR006424; GAPDH-I.
CC	-	DR	Pfam; PF00044; 9pdh; 1.
CC	-	DR	Pfam; PF02800; 9pdh; C; 1.
CC	-	DR	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	DR	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	DR	PROSITE; PS00071; GAPDH; 1.
CC	-	KW	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044; 9pdh; 1.
CC	-	FT	PFAM; PF02800; 9pdh; C; 1.
CC	-	FT	PRINTS; PR00078; G3PDHDRGNASE.
CC	-	FT	TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC	-	FT	PROSITE; PS00071; GAPDH; 1.
CC	-	FT	GLYCOLYSIS; Oxidoreductase; NAD: Multigene family; Complete proteome.
CC	-	FT	BINDING 150 150
CC	-	FT	GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
CC	-	FT	ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
CC	-	FT	PFAM; PF00044;

3	KVGINGFGRIGRLVIGRLLEVKSNIIDVVAINDLTSPIKILAYLLKHDNSNYGPF	54
[6]	SEQUENCE OF 1-13 FROM N.A.	
RP	SEQUENCE OF 1-13 FROM N.A.	
RR	STRAIN=K12; MEDLINE=88232416; PubMed=2836696;	
RA	Nakamura H., Murakami H., Yamato I., Anraku Y.;	
RT	"Nucleotide sequence of the cybb gene encoding cytochrome b561 in Escherichia coli K12.";	
RT	Mol. Gen. Genet. 212:1-5 (1988) .	
[7]	SEQUENCE OF 1-165 FROM N.A.	
RP	Krawetz S.A. ; Submitted (OCT-1990) to the EMBL/GenBank/DDBJ databases.	
[8]	SEQUENCE OF 1-52 FROM N.A.	
RP	MEDLINE=87053179; PubMed=3780374;	
RA	Krawetz S.A., Connor W., Cannon P.D., Dixon G.H.;	
RT	"A vector-primer-cloner-sequencer plasmid for the construction of cDNA libraries: evidence for a rat glyceraldehyde-3-phosphate dehydrogenase-like mRNA and a ferritin mRNA within testis.";	
RT	DNA 5:427-435 (1986) .	
-!	- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.	
CC	-!- PATHWAY: Second phase of glycolysis; first step.	
CC	-!- SUBUNIT: Homotetramer (By similarity).	
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).	
CC	-!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.	
CC	-!- CAUTION: IN THE K12 STRAIN THIS GENE IS DISRUPTED BY A STOP CODON AND A FRAMESHIFT. IT SEEMS TO BE INTACT IN A NUMBER OF WILD STRAINS.	
CC	-!- CAUTION: Ref.7 and Ref.8 sequences were originally thought to originate from rat.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	
DR	EMBL; L09067; AAA23856.1; ALT FRAME.	
DR	EMBL; AE000239; AAC74498.1; ALT FRAME.	
DR	EMBL; AE000239; AAC74499.1; ALT FRAME.	
DR	EMBL; D90780; BAA15033.1; ALT FRAME.	
DR	EMBL; D90781; BAA15038.1; ALT FRAME.	
DR	EMBL; M64541; -; NOT_ANNOTATED_CDS.	
DR	EMBL; X07569; -; NOT_ANNOTATED_CDS.	
DR	EMBL; X54798; CAA38569.1; -.	
DR	HSSP; P17721; 1HDC.	
DR	EcoGene; EG12103; gapC.	
DR	InterPro; IPR000173; GAPDH-I.	
DR	PRINTS; PR00078; G3PDHDRGNASE.	
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.	
DR	PROSITE; PS00071; GAPDH; 1.	
KW	GLYCOLYSIS; Oxidoreductase; NAD; Multigene family; Complete proteome.	
FT	FT BINDING 150	
FT	ACT_SITE 177	
FT	FT CONFLICT 39	
FT	SEQUENCE 333 AA; 35649 MW; E88223297376B0A0 CRC64;	
SQ	Query Match 41.3%; Score 940.5; DB 1; Length 333;	
SQ	Best Local Similarity 47.5%; Pred. No. 1.9e-50;	
SQ	Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;	

RL J. Bacteriol. 175:7918-7930 (1993).  
 CC -; CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
 CC -; PATHWAY: Second phase of glycolysis; first step.  
 CC -; SUBUNIT: Homotetramer (By similarity).  
 CC -; SUBCELLULAR LOCATION: Cyttoplasmic.  
 CC -; SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
 CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

CC DR U39710; AAC71523.1;  
 CC DR U02213; AAD12507.1;  
 CC DR U02178; AAD12463.1;  
 CC DR C64233; C64233.  
 CC DR P17721; 1HDG.  
 CC DR MG301;  
 CC DR InterPro; IPR000173; GAP dehydrogenase.  
 CC DR InterPro; IPR006424; GAPDH-I.  
 CC DR PF00044; spdh; 1.  
 CC DR PF02800; spdh\_C; 1.  
 CC DR PR00078; G3PDHDRGNASE.  
 CC DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
 CC DR PROSITE; PS00071; GAPDH; 1.  
 CC DR KW Glycolysis; Oxidoreductase; NAD; Complete proteome.  
 CC DR FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE  
 CC (BY SIMILARITY).  
 CC FT ACT\_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS  
 CC (BY SIMILARITY).  
 CC SQ SEQUENCE 337 AA; 37097 MW; FA1EA1966687006B CRC64;  
 CC Query Match 38.2%; Score 870.5; DB 1; Length 337;  
 CC Best Local Similarity 43.8%; Pred. No. 3.5e-46;  
 CC Matches 183; Conservative 53; Mismatches 93; Indels 89; Gaps 5;  
 CC Matches 183; Conservation 53; Mismatches 93; Indels 89; Gaps 5;  
 CC QY 30 VKVGINGFGRIGRLAARRRIQVNVEVTRINDLTDPNMLLAHLILKYDTTQGRFDGTVEVKE 89  
 CC DB 8 IRVAINGFGRIGRLVFRALLSQQRNIEIVAVNDLTHPTDLAHLKYDSAHEGF----- 59  
 CC QY 90 GGFEVNGNFIKVSAERDPENIDWATDGVEIVLEALEGTVEKDKGGFDVNGKFIKVSAEKD 149  
 CC DB 60 -----KKKVVAKLNTLMDKKRKLVLVSEKRD 84  
 CC QY 150 PEQIDWATDGVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAWIDWATDGVEIVLEAT 209  
 CC DB 85 -----PANLPWAHNIDIVVEST 102  
 CC QY 210 SFFAKKEAAEKKHLHANGAKKVVIATAPG-GNDVKTVVFNNTNHDILDGTETVSGASCTTNC 268  
 CC DB 103 GRFVSEEGASLHLQA-GAKRVIISAPARQKRTKRTVYVNVNHRKTIINAEDKTIASAASCTTNC 161  
 CC QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDILRRPARAGAAANIVPNSTGAAK 328  
 CC DB 162 LAPMVHVLEKNGFLQHGTMTVYHAYTADQRLQDAPH--SDLRRARAACNIVPNSTGAAK 219  
 CC QY 329 AIGLVVPELNGKLDGAAQRVPVPTGSVTEVTLDKNVSVDEINAAMKAASNDSEFGYTED 388  
 CC DB 220 AIGLVVPEATGKLNQMLRVPVLTGSIVELCVALEKDATEQINQAMKKAASASFRYCED 279  
 CC QY 389 PIVSSSDIVGVSYGSLFDATQTKVMEVDGSQQLVKVVSWYDNEMSYTAQLVRLTLEYFAKI 446  
 CC DB 280 EIVSSSDIVGSEHGSIFDSKLTNIEVDGNKLYKVYAWYDNESSYVNQAMKVRVVNNYCAKL 337  
 CC QY RESULT 13  
 CC ID G3P\_MYCGE STANDARD; PRT; 337 AA.  
 CC AC P47543;  
 CC DT 01-FEB-1996 (Rel. 33, Created)  
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).  
 CC GN GAPA OR GAP OR MG301.  
 CC OS Mycoplasma genitalium.  
 CC OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 CC NCBI\_TaxID=2097;  
 CC RN SEQUENCE FROM N.A.  
 CC STRAIN=ATCC 33530 / G-37;  
 CC MEDLINE=96026346; PubMed=7569993;  
 CC RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 CC RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 CC RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 CC RA Nguyen D.T., Utterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,  
 CC RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
 CC RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,  
 CC RT "The minimal gene complement of Mycoplasma genitalium.";  
 CC RL Science 270:397-403(1995).  
 CC RN SEQUENCE OF 1-81 AND 279-337 FROM N.A.  
 CC STRAIN=ATCC 33530 / G-37;  
 CC MEDLINE=94075230; PubMed=8253680;  
 CC RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;  
 CC RT "A survey of the Mycoplasma genitalium genome by using random  
 CC sequencing.";  
 CC RESULT 14  
 CC ID G3P\_CORGL  
 CC DB 280 PIVSSSDIVGVSYGSLFDATQTKVMEVDGSQQLVKVVSWYDNEMSYTAQLVRLTLEYFAKI 446  
 CC RN SEQUENCE OF 1-81 AND 279-337 FROM N.A.  
 CC STRAIN=ATCC 33530 / G-37;  
 CC MEDLINE=94075230; PubMed=8253680;  
 CC RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;  
 CC RT "A survey of the Mycoplasma genitalium genome by using random  
 CC sequencing.";



RA Tosato V., Uchiyama S., Vandembol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenerger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshihikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis";  
 RL Nature 390:249-256 (1997).  
 RN [3]  
 RP SEQUENCE OF 1-30.  
 RC STRAIN=168 / JH642;  
 RX MEDLINE=96345629; PubMed=8755892;  
 RA Graumann P., Schroeder K., Schmid R., Marahiel M.A.;  
 RT "Cold shock stress-induced proteins in Bacillus subtilis.";  
 RL J. Bacteriol. 178:4611-4619(1996).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=20261518; PubMed=10799476;  
 RA Fillinger S., Boschi-Muller S., Azza S., Dervyn E., Branlant G.,  
 RA Aymerich S.;  
 RT "Two glyceraldehyde-3-phosphate dehydrogenases with opposite  
 physiological roles in a nonphotosynthetic bacterium.";  
 RL J. Biol. Chem. 275:14031-14037 (2000);  
 CC !- FUNCTION: More active in catabolism.  
 CC !- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
 CC !- PATHWAY: Second phase of glycolysis; first step.  
 CC !- SUBUNIT: Homotetramer.  
 CC !- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC !- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
 CC dehydrogenase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; X13011; CAA31434.1; -.  
 DR EMBL; Z99121; CAB15399.1; -.  
 DR PIR; S02754; DEBSG.  
 DR HSSP; P00362; 1GD1.  
 DR SubtilList; BG108227; gapA.  
 DR InterPro; IPR00173; GAP dhydrogenase.  
 DR InterPro; IPR006424; GAPDH-I.  
 DR Pfam; PF00044; gpdh.  
 DR Pfam; PF02800; gpdh\_C; 1.  
 DR PRINTS; PR00078; G3PDHDRGNASE.  
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
 DR PROSITE; PS00071; GAPDH; 1.  
 KW GLYCOLYSIS; Oxidoreductase; NAD; Multigene family; Complete proteome.  
 FT INIT MET 0 0  
 FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE.  
 FT ACT SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.  
 SQ SEQUENCE 334 AA; 35701 MW; 1283D3E6CF5095EC CRC64;  
 Query Match 37.1%; Score 844; DB 1; Length 334;  
 Best Local Similarity 43.7%; Fred. No. 1.4e-44;  
 Matches 184; Conservative 45; Mismatches 96; Indels 96; Gaps 7;  
 QY 30 VKVGINFGGRIGRLLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTVEVKE 89  
 Db 2 VKVGINFGGRIGRNVRAALNNPEVEVAVNDLTDAANMLAHLQQYDSVHGLDAEVSV-- 59  
 QY 90 GGFEVNGNEFIKVSSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAEKD 149  
 Db 60 -----  
 QY 150 PEQIDWATDGVEIVLELEIDGTVEVKEGGFEVNGQFVKVSAEREPAVIDWATDGVEIVLEAT 209  
 Db 60 -----DGNNLV-----VNGKTIEVSAERDPAKLSWGLQGVEIVVEST 96

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:18:54 ; Search time 119 Seconds  
 (without alignments)  
 1187.833 Million cell updates/sec

Title: US-10-650-369-22  
 Perfect score: 2278  
 Sequence: 1 MKKRTGILLLAVIILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing First 45 summaries

Database : SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rat:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriopl:\*
- 17: sp\_archeap:\*

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rat:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriopl:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1656.5	72.7	336	2	Q83ZF7		Q83zf7 streptococc
2	1564.5	68.7	336	2	Q9AIW2		Q9AIW2 streptococc
3	1564.5	68.7	336	16	Q8E3E8		Q8E3E8 streptococc
4	1564.5	68.7	336	16	Q8DXS8		Q8DXS8 streptococc
5	1559.5	68.5	336	2	Q8KVU6		Q8KVU6 streptococc
6	1544.5	67.8	336	2	Q8GCR7		Q8GCR7 streptococc
7	1535	67.4	335	16	Q97NL1		Q97NL1 streptococc
8	1535	67.4	359	16	Q8CWN6		Q8CWN6 streptococc
9	1534.5	67.4	336	2	Q8KHG1		Q8KHG1 streptococc
10	1530.5	67.2	336	2	Q8KVU5		Q8KVU5 streptococc
11	1521	66.8	337	16	Q8DVV3		Q8DVV3 streptococc
12	1513.5	66.4	336	2	Q8VVB9		Q8VVB9 streptococc
13	1459.5	64.1	320	2	Q9L5X6		Q9L5X6 streptococc
14	1410	61.9	309	2	Q9AJT7		Q9AJT7 streptococc
15	1409	61.9	309	2	Q9AJT9		Q9AJT9 streptococc
16	1407	61.8	309	2	Q9AJT4		Q9AJT4 streptococc

Query Match 72.7%; Score 1656.5; DB 2; Length 336;  
 Best Local Similarity 79.6%; Pred. No. 2.3e-88;  
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;  
 QY 28 MVVKVINGFGRIGRILAFRRRIQNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVY 87  
 DB 1 MVVKVINGFGRIGRILAFRRRIQNVEGVEVTRINDLTDPMMLAHLKYDTTQGRFDGTVY 60

[1] SEQUENCE FROM N.A.  
 RA Perez-Casal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;  
 RT "Use of the surface proteins GapC and Mig of Streptococcus  
 RT dysgalactiae as protective antigens against mastitis in non-lactating  
 RT cows";  
 RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF375662; AAP31408.1;  
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.  
 DR GO; GO:0006096; P:glycolysis; IEA.  
 DR InterPro; IPR000173; GAP\_dhdrogenase.  
 DR Pfam; PF0044; Gpdh; 1.  
 DR Pfam; PF02800; Gpdh; C; 1.  
 DR PRINTS; PR00078; G3PDHDRGNASE.  
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.  
 DR PROSITE; PS00071; GAPDH; 1.  
 SQ SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;

QY	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLAELEGTVKDGFDVNGKFIKVSAAE	147	QY	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGVEIVLAELEGTVKDGFDVNGKFIKVSAAE	147
Db	61	KEGGFEVNGNFIKVSAAERDP-----	81	Db	54	-----	54
QY	148	KDPEQIDWATDGVEIVLIEDGTVEVKEGGFVNQFVKAEREPANIDWATDGVEIVL-----	207	QY	148	KDPEQIDWATDGVEIVLIEDGTVEVKEGGFVNQFVKAEREPANIDWATDGVEIVL-----	207
Db	82	-----	-----	Db	55	-----	55
QY	208	ATSEFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVIGASCTTN	267	QY	208	ATSEFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVIGASCTTN	267
Db	96	ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVIGASCTTN	155	Db	96	ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNNTNHDILDGTETVIGASCTTN	155
QY	268	CLAPMAKALHDAFGIQLKGLMTTIHAYTGQMDLGPGRGGDLRRARAAGANTVNPNSTGAA	327	QY	268	CLAPMAKALHDAFGIQLKGLMTTIHAYTGQMDLGPGRGGDLRRARAAGANTVNPNSTGAA	327
Db	156	CLAPMAKALHDAFGIQLKGLMTTIHAYTGQMDLGPGRGGDLRRARAAGANTVNPNSTGAA	215	Db	156	CLAPMAKALQDNFGVKGQLGLMTTIHAYTGQMDLGPGRGGDLRRARAAGANTVNPNSTGAA	215
QY	328	KAIGLVIPELNGKLQDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGY	387	QY	328	KAIGLVIPELNGKLQDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGY	387
Db	216	KAIGLVIPELNGKLQDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGY	275	Db	216	KAIGLVIPELNGKLQDGAAQRVPVPTGSVTELVVTLDKNVSVDEINAAMKAASNDSEFGY	275
QY	388	DPIVSSSDIVGVSYGSLSFDATQTKVMEVTDGSQQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA	447	QY	388	DPIVSSSDIVGVSYGSLSFDATQTKVMEVTDGSQQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA	447
Db	276	DPIVSSSDIVGVSYGSLSFDATQTKVMEVTDGSQQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA	335	Db	276	DPIVSSSDIVGVSYGSLSFDATQTKVMEVTDGSQQLVKKVSVWYDNEMSYTAQLVRLTLEYFAKIA	335
QY	448	K 448		QY	448	K 448	
Db	336	K 336		Db	336	K 336	
RESULT 2							
Q9ALW2		PRELIMINARY;	PRT;	336	AA.		
ID	Q9ALW2;						
AC	Q9ALW2;						
DT	01-JUN-2001	(TREMBLrel.	17, Created)				
DT	01-JUN-2001	(TREMBLrel.	17, Last sequence update)				
DT	01-OCT-2003	(TREMBLrel.	25, Last annotation update)				
DE							
OS							
OC							
RN							
RP	SEQUENCE FROM N.A.						
RC	STRAIN=J48;						
RA	Seifert K.N., Bleiweis A.S., McArthur W.P., Brady L.J.;						
RT	"The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde 3-Phosphate Dehydrogenase."						
RL	Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.						
CC	-!!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.						
DR	EMBL; AF338416; AAK14387.1; -.						
DR	HSSP; P00362; 1GDI.						
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase activity; IEA.						
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.						
DR	GO; GO:0006096; P:glycolysis; IEA.						
DR	InterPro; IPR006424; GAPDH-I.						
DR	InterPro; IPR000173; GAP_dhdrogenase.						
DR	PFam; PF00044; gpdh; 1.						
DR	PFam; PF02800; gpdh_C; 1.						
DR	PRINTS; PR00078; G3PDHGNASE.						
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.						
DR	PROSITE; PS00071; GAPDH; 1.						
KW	Oxidoreductase.						
SEQUENCE	336 AA;	36005 MW;	982E18A05CA343C9 CRC64;				
SQ							
Query Match	68.7%;	Score 1564.5;	DB 2;	Length 336;			
Best Local Similarity	74.1%;	Pred. No. 5e-83;					
Matches	312;	Conservative 15;	Mismatches 9;	Indels 85;	Gaps 1;		
QY	28	MVVKVGINGFGRIGRLAERRRIQNEVGEVTRINDLTDPNMLAHLKYDTTQGRFDGTV	87				
Db	1	MVVKVGINGFGRIGRLAERRRIQNEVGEVTRINDLTDPNMLAHLKYDTTQGRFDGTV	54				
Query Match	68.7%;	Score 1564.5;	DB 16;	Length 336;			
Best Local Similarity	74.1%;	Pred. No. 5e-83;					
Matches	312;	Conservative 15;	Mismatches 9;	Indels 85;	Gaps 1;		
QY	28	MVVKVGINGFGRIGRLAERRRIQNEVGEVTRINDLTDPNMLAHLKYDTTQGRFDGTV	87				

Db	1	MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRF-----	54	Query Match 68.7%; Score 1564.5; DB 16; Length 336;
Qy	88	KEGGFEVNGNFIKVSAAERDPENIDWATDGEIVLEALEGGTEVKDGGFDVNGKFIKVSAAE	147	Best Local Similarity 74.1%; Pred. No. 5e-83; Mismatches 15; Conservative 15; Indels 85; Gaps 1;
Db	55	-	-	Matches 312; Conservative 15; Indels 85; Gaps 1;
Qy	148	KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPA	207	Qy 28 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRF-----
Db	55	-	-	Db 1 MVVKVGINGFGRIGRLAFRRIONVEGVEVTRINDLDPNMLAHLKYDTTQGRF-----
Qy	208	ATSFFAKKAAEKHLHANGAKKVITAPGGNDVKTVFFNTNHDILDGTETVSGASCTTN	267	Qy 54 88 KEGGFEVNGNFIKVSAAERDPENIDWATDGEIVLEALEGGTEVKDGGFDVNGKFIKVSAAE
Db	96	ATGF FASKEKAEQHIIHENGAKKVITAPGGNDVKTVFFNTNHDILDGTETVSGASCTTN	155	Db 55 -
Qy	268	CLAPMAKALHDAFGIQLGMLTTIHAYTGQMDMILDGPHRGDDLRARRAGAAANIVPNSTGAA	327	Qy 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPA
Db	156	CLAPMAKALQDNFGVKQGLMLTTIHAYTGQMDMILDGPHRGDDLRARRAGAAANIVPNSTGAA	215	Db 55 -
Qy	328	KAIGLVIPELNGKLDGAQQRVPVPTGSVTELVYVTLDKNVSVDEINAAMKAASNDSEGYTE	387	Qy 208 ATSF FAKKAAEKHLHANGAKKVITAPGGNDVKTVFFNTNHDILDGTETVSGASCTTN
Db	216	KAIGLVIPELNGKLDGAQQRVPVPTGSVTELVATLEKDVTVEVNAAMKAANDSYGYTE	275	Db 96 ATGFFASKEKAEQHIIHENGAKKVITAPGGNDVKTVFFNTNHDILDGTETVSGASCTTN
Qy	388	DPIVSSDIVGVSYGSLFDATOTKVMEVVDGSQLVKVVSWYDNEMSYTAQLVRLTEYFAKIA	447	Qy 268 CLAPMAKALHDAFPQIQLGMLTTIHAYTGQMDMILDGPHRGDDLRARRAGAAANIVPNSTGAA
Db	276	DPIVSSDIVGVSYGSLFDATOTKVQTVDGNQLVKVVSWYDNEMSYTSQQLVRLTEYFAKIA	335	Db 156 CLAPMAKALQDNFGVKQGLMLTTIHAYTGQMDMILDGPHRGDDLRARRAGAAANIVPNSTGAA
Qy	448	K 448	Qy 216 KAIGLVIPELNGKLDGAQQRVPVPTGSVTELVATLEKDVTVEVNAAMKAANDSYGYTE	
Db	336	K 336	Qy 276 DPIVSSDIVGVSYGSLFDATOTKVQTVDGNQLVKVVSWYDNEMSYTSQQLVRLTEYFAKIA	
<b>RESULT 4</b>				
Q8DXS8		PRELIMINARY; PRT; 336 AA.	QY 336 K 448	
ID			Db 336 K 336	
AC				
Q8DXS8;				
DT	01-MAR-2003	(TREMBLrel. 23, Created)		
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE	Glyceraldehyde 3-phosphate dehydrogenase.			
GN	GAP OR SAG1768.			
OS	Streptococcus agalactiae (serotype V).			
OC	Bacteria; Firmicutes; Streptococceae; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=216466;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=2603 V/R / Serotype V;			
RX	MEDLINE=22222988; PubMed=12200547;			
RA	Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,			
RA	Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,			
RA	Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,			
RA	DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,			
RA	Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,			
RA	Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,			
RA	Iacobini E.T., Brettoni C., Galli G., Mariani M., Vigni F., Maione D.,			
RA	Rinaudo D., Rappuoli R., Telford J.L., Kaspar D.L., Grandi G.,			
RA	Fraser C.M.;			
RT	"Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).			
DR	EMBL; AE014272; AAN00631.1; -.			
DR	TIGR; SAG1768; -.			
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.			
DR	GO; GO:0006096; P:glycolysis; IEA.			
DR	InterPro; IPR006424; GAPDH-I.			
DR	InterPro; IPR000173; GAP_dhydrogenase.			
DR	Pfam; PF00044; gpdh; 1.			
DR	PFam; PF02800; gpdh_C; 1.			
DR	PRINTS; PR00078; G3PDHDRGNASE.			
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.			
DR	PROSITE; PS00071; GAPDH; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;			





RP	SEQUENCE FROM N.A.	OX	NCBI_TAXID=1348;
RC	Fontaine S.agalactiae, and S.iniae; STRAIN=ATCC 27541, and 9117;	RN	[1] _TAXID=1348;
RA	Potter A.A.; Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,	RP	SEQUENCE FROM N.A.
RA	"Immunization of dairy cattle with recombinant GapC and chimeric CAMP	RA	Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RT	antigens confers protection against heterologous challenge with	RA	Potter A.A.;
RT	Streptococcus uberis.";	RT	"Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.	RT	antigens confers protection against heterologous challenge with
RL	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE	RT	Streptococcus uberis.";
CC	DEHYDROGENASE FAMILY.	RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC	DEHYDROGENASE FAMILY.	CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
DR	EMBL; AF421899; AAM73770.1; -.	CC	DEHYDROGENASE FAMILY.
DR	EMBL; AF421901; AAM73773.1; -.	DR	DR; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.	DR	DR; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0006096; P:glycolysis; IEA.	DR	DR; GO:0006096; P:glycolysis; IEA.
DR	InterPro; IPR006424; GAPDH-I.	DR	InterPro; IPR006424; GAPDH-I.
DR	InterPro; IPR000173; GAP_dhydrogenase.	DR	InterPro; IPR000173; GAP_dhydrogenase.
DR	PFam; PF00044; gpdh_1.	DR	PFam; PF00044; gpdh_1.
DR	PFam; PF02800; gpdh_C; 1.	DR	PFam; PF02800; gpdh_C; 1.
DR	PRINTS; PR00078; G3PDHDRGNASE.	DR	PRINTS; PR00078; G3PDHDRGNASE.
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.	DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR	PROSITE; PS00071; GAPDH; 1.	DR	PROSITE; PS00071; GAPDH; 1.
KW	Oxidoreductase.	KW	Oxidoreductase.
SQ	SEQUENCE 336 AA; 35723 MW; ABAB9E14F3ED1111 CRC64;	SQ	SEQUENCE 336 AA; 36037 MW; 311647C225489AC9E CRC64;
Query	Query Match 67.4%; Score 1534.5; DB 2; Length 336;	Query	Query Match 67.2%; Score 1530.5; DB 2; Length 336;
Best Local Similarity 72.7%; Pred. No. 2.7e-81;	Best Local Similarity 73.4%; Pred. No. 4.6e-81;	Best Local Similarity 73.4%; Pred. No. 4.6e-81;	Best Local Similarity 73.4%; Pred. No. 4.6e-81;
Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;	Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;	Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;	Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;
QY	28 MVVKVINGFGRIGRLAFRRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTVEV 87	QY	28 MVVKVINGFGRIGRLAFRRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTVEV 87
Db	1 MVVKVINGFGRIGRLAFRRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFDGTVEV 60	Db	1 MVVKVINGFGRIGRLAFRRIQNVEGVETRINDLTDPMMLAHLKYDTTQGRFD--- 55
QY	88 KEGGFEVNGNF1KVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAE 147	QY	88 KEGGFEVNGNF1KVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAE 147
Db	61 KDGGEFVNNGS-----70	Db	56 -----GTVEVKDGGFDVNGKFIKVSAE 77
QY	148 KDPEQIDWATDGVEIVLEIDGTVEVKEGGFREVNGQFVVKVSAEREPANIDWATDGVEIVLE 207	QY	148 KDPEQIDWATDGVEIVLEIDGTVEVKEGGFREVNGQFVVKVSAEREPANIDWATDGVEIVLE 207
Db	71 -----FVKVSAEREPANIDWATDGVDIVLE 95	Db	78 KDPEQ-----IDWATDGVEIVLE 95
QY	208 ATSEFFAKKEAAEKLHLHANGAKKVVITAPGGNDVKTVVFTNHDILDGTETVSGASCTTN 267	QY	208 ATSEFFAKKEAAEKLHLHANGAKKVVITAPGGNDVKTVVFTNHDILDGTETVSGASCTTN 267
Db	96 ATGFFASKAAAEEQHITHANGAKKVVITAPGGNDVKTVVFTNHDILDGTETVSGASCTTN 155	Db	96 ATGFFAKKEAAEKLHLHENGAKKVVITAPGGNDVKTVVFTNHDILDGTETVSGASCTTN 155
QY	268 CLAPMAKALHDAGIQLQGLMTTIHAYTGDQMILDGPHRGGLDLRARAAGAANIVPNSTGAA 327	QY	268 CLAPMAKALHDAGIQLQGLMTTIHAYTGDQMILDGPHRGGLDLRARAAGAANIVPNSTGAA 327
Db	156 CLAPMAKALQDNFGVKQGLMTTIHGTGDQMVLDDGPHRGGLDLRARAAGAANIVPNSTGAA 215	Db	156 CLAPMAKALQDNFGVKQGLMTTIHAYTGDQMILDGPHRGGLDLRARAAGAANIVPNSTGAA 215
QY	328 KAIGLVIPELINGKLDGAAQRVYPVPTGSVTELVVTLDKNVSVDINAAMKAASNDSFGYTE 387	QY	328 KAIGLVIPELINGKLDGAAQRVYPVPTGSVTELVVTLDKNVSVDINAAMKAASNDSFGYTE 387
Db	216 KAIGLVIPELINGKLDGAAQRVYPVPTGSVTELVAVLEKDTSEEVINAAMKAASNDSFGYTE 275	Db	216 KAIGLVIPELINGKLDGAAQRVYPVPTGSVTELVAVLNKETSSEEINSMKAASNDSFGYTE 275
QY	388 DPIVSSDIVGYSYGSLSFDATQTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRTLEYFAKIA 447	QY	388 DPIVSSDIVGYSYGSLSFDATQTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db	276 DAIVSSDIVGYSYGSLSFDATQTKVQTVDGNQLVKVTVSWYDNEMSYTAQLVRTLEYFAKIA 335	Db	276 DAIVSSDIVGYSYGSLSFDATQTKVQTVDGNQLVKVTVSWYDNEMSYTAQLVRTLEYFAKIA 335
QY	448 K 448	QY	448 K 448
Db	336 K 336	Db	336 K 336
RESULT 10	RESULT 11		
Q8KVU5	PRELIMINARY; PRT; 336 AA.	Q8DVV3	PRELIMINARY; PRT; 337 AA.
ID		ID	
AC		AC	
DT	01-OCT-2002 (TREMBLrel. 22, Created)	DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)	DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Glyceraldehyde-3-phosphate dehydrogenase.	DE	Extracellular glyceraldehyde-3-phosphate dehydrogenase (EC
OS	Streptococcus parauberis.	OS	1.2.1.12).
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	GN	GAPC OR SMU 360.
OC	Streptococcus mutans.	OS	

OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	OS	Streptococcus thermophilus.
OC	Streptococcus.	OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX	NCBI_TaxID=1309;	OC	Streptococcus.
RN	[1]	OX	NCBI_TaxID=1308;
RP	SEQUENCE FROM N.A.	RN	RN
RC	STRAIN=UA159 / ATCC 700610 / Serotype C;	RP	SEQUENCE FROM N.A.
RX	MEDLINE=222295063; PubMed=12397186;	RC	STRAIN=LMG118311;
RA	Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,	RA	van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA	Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S.,	RA	Kuipers O.P., de Vos W.M.,
RA	Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;	RA	"Modulation of Glycolysis by Lactose Availability in Streptococcus thermophilus";
RT	"Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen.";	RT	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RT	RT -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY	RT	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY
RL	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).	CC	CC
DR	EMBL; AE014883; AAN58118.1; -.	DR	EMBL; AF442551; AAL35377.1; -.
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.	DR	HSSP; P00354; 3GPD.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.	DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR	GO; GO:006096; P:glycolysis; IEA.	DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	InterPro; IPR006424; GAPDH-I.	DR	GO; GO:0006096; P:glycolysis; IEA.
DR	InterPro; IPR000173; GAP_dhydrogenase.	DR	InterPro; IPR000173; GAP_dhydrogenase.
DR	Pfam; PF00044; gpdh; 1.	DR	InterPro; IPR000173; GAP_dhydrogenase.
DR	Pfam; PF02800; gpdh; C; 1.	DR	Pfam; PF00044; gpdh; 1.
DR	PRINTS; PR00078; G3PDHDRGNASE.	DR	PRINTS; PR00078; G3PDHDRGNASE.
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.	DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR	PROSITE; PS00071; GAPDH; 1.	DR	PROSITE; PS00071; GAPDH; 1.
KW	Oxidoreductase; Complete proteome.	KW	Oxidoreductase; Complete proteome.
SQ	SEQUENCE 337 AA; 36068 MW; 42BFE20365963C22 CRC64;	SQ	SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;
Query Match	66.8%; Score 1521; DB 16; Length 337;	Query Match	66.4%; Score 1513.5; DB 2; Length 336;
Best Local Similarity	71.8%; Pred. No. 1.7e-80;	Best Local Similarity	71.7%; Pred. No. 4.5e-80;
Matches	303; Conservative 18; Mismatches 15; Indels 86; Gaps 2;	Matches	302; Conservative 18; Mismatches 16; Indels 85; Gaps 1;
QY	28 MVVKVGINGFGRIGLAFRRIONVNEGVETRINDLDPNMLLAHLLKYDTTQGRFDTGTEV 87	QY	28 MVVKVGINGFGRIGLAFRRIONVNEGVETRINDLDPNMLLAHLLKYDTTQGRFDTGTEV 87
Db	1 MVVKVGINGFGRIGLAFRRIONVNEGVETRINDLDPNMLLAHLLKYDTTQGRFDTGTEV 60	Db	1 MVVKVGINGFGRIGLAFRRIONVNEGVETRINDLDPNMLLAHLLKYDTTQGRFDTGTEV 60
QY	88 KEGGFEVNGNF1KVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFKVSAE 147	QY	88 KEGGFEVNGNF1KVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFKVSAE 147
Db	61 KEGGFEVNGKFVFKVSAERDP-----	Db	61 KEGGFEVNGKFVFKVSAERDP-----
QY	148 KDPEQIDWATDGVEIVLEIDGTVEVKDGGFEVNGQFVKVSAEREPANIDWATDGVEIVLE 207	QY	148 KDPEQIDWATDGVEIVLEIDGTVEVKDGGFEVNGQFVKVSAEREPANIDWATDGVEIVLE 207
Db	82 -----	Db	82 -----
QY	208 ATSFPAKEAAEKHLHANAKKVVITAPGGNDVFKTVFNTNHDILDGTETVISGASCTT 266	QY	208 ATSFPAKEAAEKHLHANAKKVVITAPGGNDVFKTVFNTNHDILDGTETVISGASCTT 267
Db	96 ATGFFASKAAAEKHLHANGGAKKVVITAPGGNDIKTIVFNTNHDILDGTETVISGASCTT 155	Db	96 ATGFFFTKKVLAEKHLHPGGAKKVVITAPGGNDVFKTVFNTNHDILDGTETVISGASCTT 155
QY	267 NCLAPMAKALHDAFGIQKGLMTTHIAYTGDQMILDGPHRGGDLRRARAGAANIVPNSTGA 326	QY	268 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQMILDGPHRGGDLRRARAGAANIVPNSTGA 327
Db	156 NCLAPMAKALHDNFSIKEGLMTTHIAYTGDQMILDGPHRGGDLRRARAGAANIVPNSTGA 215	Db	156 CLAPMAKALNDNFGIVEGLGLMTTHIAYTGDQMILDGPNRGGLDRLRARAGAANIVPNSTGA 215
QY	327 AKAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSDEINAAMKAASNDSFGYT 386	QY	328 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNVSDEINAAMKAASNDSFGYT 387
Db	216 AKAIGLVIPELNGKLDGAAQRVPVPTGSVTELVAVLDDKKVTVDEVNAAMKAANESGYT 275	Db	216 KAIGLVIPELNGKLDGAAQRVPVPTGSVTELVAVCBEKXNTVDEVNAAMKAATNESGYT 275
QY	387 EDPIVSSDIVGVSYGSLSFDATAQTKVMEVDGSQLVKKVWSYDNEMSYTAQLVRLTLEYFAK 446	QY	388 DPIVSSDIVGVSYGSLSFDATAQTKVMEVDGSQLVKKVWSYDNEMSYTAQLVRLTLEYFAK 447
Db	276 EDPIVSSDIVGMSFGSLFDATAQTRVLDVDGKQLVKKVWSYDNEMSYTSQIVRLTLEYFAK 335	Db	276 EDPIVSSDIVGMSFGSLFDATAQTRVLDVDGKQLVKKVWSYDNEMSYTSQIVRLTLEYFAK 335
QY	447 AK 448	QY	448 K 448
Db	336 AK 337	Db	336 K 336
RESULT 12	Q8VVVB9	RESULT 13	Q9L5X6
ID	Q8VVVB9	PRELIMINARY;	PRELIMINARY;
AC	Q8VVVB9;	PRT;	PRT;
DT	01-MAR-2002 (TREMBLrel. 20, Created)	DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Glyceraldehyde-3-phosphate dehydrogenase.	GN	GAPDH.

DE	Extracellular glyceraldehyde-3-phosphate dehydrogenase	DB	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
OS	(Fragment).	DE	(Fragment).
OC	Streptococcus gordonii.	GN	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	OS	Streptococcus pneumoniae.
OX	Streptococcus	OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
NCBI_TaxID=1302;	NCBI_TaxID=1313;	OC	Streptococcus.
RN	[1]	RN	RN
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RA	Nelson D., Goldstein J.M., Boatright K., Harry D.W.S., Cook S.L., Hickman P.J., Potempa J., Travis J., Mayo J.A.;	RA	SEQUENCE FROM N.A.
RA	"Purification and characterization of an extracellular glyceraldehyde-3-phosphate dehydrogenase from Streptococcus sanguis and cloning of the gene encoding this enzyme"; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	RA	SEQUENCE FROM N.A.
RT	"Sequencing variation in <i>mef</i> (A) and <i>gap</i> genes in M-phenotype erythromycin resistant Streptococcus pneumoniae.";	RT	Amezaga M.R., Carter P.E., Cash P., McKenzie H.;
RT	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.	RT	"Sequencing variation in <i>mef</i> (A) and <i>gap</i> genes in M-phenotype erythromycin resistant Streptococcus pneumoniae.";
RL	CC	RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.	CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.
DR	EMBL; AF247678; AAF64063.1; -.	DR	EMBL; AJ292048; CAC27448.1; -.
DR	HSSP; P00362; 1GD1.	DR	HSSP; P00362; 1GD1.
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.	DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.	DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0006096; P:glycolysis; IEA.	DR	GO; GO:0006096; P:glycolysis; IEA.
DR	InterPro; IPR006424; GAPDH-I.	DR	InterPro; IPR006424; GAPDH-I.
DR	PFam; PF00044; gpdh; 1.	DR	PFam; PF00044; gpdh; 1.
DR	PFam; PF02800; gpdh_C; 1.	DR	PFam; PF02800; gpdh_C; 1.
DR	PRINTS; PR00078; G3PDHDRGNASE.	DR	PRINTS; PR00078; G3PDHDRGNASE.
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.	DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR	PROSITE; PS00071; GAPDH; 1.	DR	PROSITE; PS00071; GAPDH; 1.
KW	Oxidoreductase.	KW	Oxidoreductase.
FT	NON-TER	FT	NON-TER
FT	NON-TER	FT	NON-TER
SQ	SEQUENCE 320 AA; 320 MW; 33968 MW; 6FFCC18BB01E91F0 CRC64;	SQ	SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;
Query Match	64.1%; Score 1459.5; DB 2; Length 320;	Query Match	61.9%; Score 1410; DB 2; Length 309;
Best Local Similarity	71.9%; Pred. No. 5.7e-77; Mismatches 14; Indels 85; Gaps 1;	Best Local Similarity	71.9%; Pred. No. 4e-74; Mismatches 13; Indels 86; Gaps 2;
Matches	291; Conservative	Matches	284; Conservative
Qy	31 KVGINGFGRIGRLAFRRIONVEGVTRINDLTDPMMLAHLKYDTTQGRFDGTVEK 90	Qy	36 GFGRIGRLAFRRIONVEGVTRINDLTDPMMLAHLKYDTTQGRFDGTVEK 95
Db	1 KVGINGFGRIGRLAFRRIONVEGVTRINDLTDPMMLAHLKYDTTQGRFDGTVEK 60	Db	1 GFGRIGRLAFRRIONVEGVTRINDLTDPMMLAHLKYDTTQGRFDGTVEK 60
Qy	91 GFEVNGNFIKVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAEKDP 150	Qy	96 GNF1KVSAERDPENIDWATDGVEIVLEALEGTVEVKDGGFDVNGKFIKVSAEKDPQIDW 155
Db	61 GFEVNGKFKVKAERDP--	Db	61 GKFIKVSAERDP--
Qy	151 EQIDWATDGVEIVLEIDGTVEVKEGGEVNGQFVKVSAEREPA	Qy	156 ATDGVEIVLEIDGTVEVKEGGEVNGQFVKVSAEREPA
Db	79 -----	Db	74 -----
Qy	211 FFAKKEAAEKLHANGAKKVITAPGGNDVKTWVENTNHDILDGTETVSGASCTTNCLA 270	Qy	216 EAAEKHLHANGAKKVITAPGGNDVKTWVENTNHDILDGTETVSGASCTTNCLA 275
Db	96 FFATKAAAEKHLHAGGAKKVITAPGGSDVKTWVENTNHDILDGTETVSGASCTTNCLA 155	Db	96 EAAEKHL-KGGAKKVVITAPGGNDVKTWVENTNHDILDGTETVSGASCTTNCLA 154
Qy	271 PMAKALHDAGF1QKGLMTTIHAYTGDQMILDGPHRGGLDRLRARAAGANTVNSTGAAKAI 330	Qy	276 LHDAFGICKGLMTTIHAYTGDQMILDGPHRGGLDRLRARAAGANTVNSTGAAKAI 335
Db	156 PMAKALQDNFGVVEGLMTTIHAYTGDQMILDGPHRGGLDRLRARAAGANTVNSTGAAKAI 215	Db	155 LQDNFGVVEGLMTTIHAYTGDQMILDGPHRGGLDRLRARAAGANTVNSTGAAKAI 214
Qy	331 GLVIPELNGKLDGAAQRVPVPTGSVTELVVTLDKNTVSDEINAAMKAASNSDFGYTEDPI 390	Qy	336 ELNGKLDGAAQRVPVPTGSVTELVVTLDKNTVSDEINAAMKAASNSDFGYTEDPI 395
Db	216 GLVIPELNGKLDGAAQRVPVPTGSVTELVAVLDKNTVDEVNAAAMKAASNSESGYTEDPI 275	Db	215 ELNGKLDGSAQRVPVPTGSVTELVAVLDKNTVDEVNAAAMKAASNSESGYTEDPI 274
Qy	391 VSSDIVGVSYGSLFDATQKVMEVDGSQLVQVSWYDNEMSYTAQ 435	Qy	396 VGVSYGSLSFDATQKVMEVDGSQLVQVSWYDNEM 430
Db	276 VSSDIVGMAYGSLFDATQTKVLDVDGKQLVKVSWYDNEMSYTAQ 320	Db	275 VGMGSYGSLSFDATQKVLVDGKQLVKVSWYDNEM 309
RESULT	14	RESULT	15
Q9AJT7	PRELIMINARY; PRT; 309 AA.	Q9AJT9	PRELIMINARY; PRT; 309 AA.
ID	Q9AJT7	ID	Q9AJT9
AC	Q9AJT7;	AC	Q9AJT9;
DT	01-JUN-2001 (TREMBLrel. 17, Created)	DT	01-JUN-2001 (TREMBLrel. 17, Created)
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)



This Page Blank (uspto)